

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 14:43:11 ; Search time 18.33 Seconds
(without alignments)
424.776 Million cell updates/sec

Title: US-09-407-806A-4
Perfect score: 346
Sequence: 1 LRALVFHGNLQYAEIPKSEP.....ERRLDAFRAIYNDWRNGEP 346

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 346 | 100.0 | 346 | 2 | US-08-613-220B-4 |
| 2 | 8 | 2.3 | 466 | 4 | US-09-068-195-25 |
| 3 | 7 | 2.0 | 113 | 4 | US-08-905-223-439 |
| 4 | 7 | 2.0 | 301 | 3 | US-09-002-298-3 |
| 5 | 7 | 2.0 | 484 | 4 | US-08-913-578-2 |
| 6 | 7 | 2.0 | 484 | 4 | US-08-785-427-2 |
| 7 | 7 | 2.0 | 549 | 2 | US-08-676-279-59 |
| 8 | 6 | 1.7 | 11 | 4 | US-09-196-293-7 |
| 9 | 6 | 1.7 | 11 | 4 | US-08-209-603E-7 |
| 10 | 6 | 1.7 | 13 | 5 | PCT-US94-10257A-52 |
| 11 | 6 | 1.7 | 15 | 1 | US-08-378-761A-48 |
| 12 | 6 | 1.7 | 15 | 1 | US-08-485-286-48 |
| 13 | 6 | 1.7 | 15 | 6 | 5248606-29 |
| 14 | 6 | 1.7 | 18 | 1 | US-08-786-748A-128 |
| 15 | 6 | 1.7 | 18 | 1 | US-08-786-748A-132 |
| 16 | 6 | 1.7 | 18 | 1 | US-08-786-748A-134 |
| 17 | 6 | 1.7 | 18 | 2 | US-08-932-682-128 |
| 18 | 6 | 1.7 | 18 | 2 | US-08-932-682-132 |
| 19 | 6 | 1.7 | 18 | 2 | US-08-932-682-134 |
| 20 | 6 | 1.7 | 20 | 5 | PCT-US94-10257A-80 |
| 21 | 6 | 1.7 | 21 | 1 | US-08-786-748A-123 |
| 22 | 6 | 1.7 | 21 | 1 | US-08-786-748A-126 |
| 23 | 6 | 1.7 | 21 | 1 | US-08-786-748A-127 |
| 24 | 6 | 1.7 | 21 | 1 | US-08-786-748A-135 |
| 25 | 6 | 1.7 | 21 | 1 | US-08-786-748A-137 |
| 26 | 6 | 1.7 | 21 | 2 | US-08-932-682-123 |
| 27 | 6 | 1.7 | 21 | 2 | US-08-932-682-126 |

| | | | | | | |
|----|---|-----|-----|---|-------------------|--------------------|
| 28 | 6 | 1.7 | 21 | 2 | US-08-932-682-127 | Sequence 127, App |
| 29 | 6 | 1.7 | 21 | 2 | US-08-932-682-135 | Sequence 135, App |
| 30 | 6 | 1.7 | 21 | 2 | US-08-932-682-137 | Sequence 137, App |
| 31 | 6 | 1.7 | 67 | 4 | US-08-988-856B-17 | Sequence 17, Appl |
| 32 | 6 | 1.7 | 95 | 2 | US-08-461-990B-26 | Sequence 26, Appl |
| 33 | 6 | 1.7 | 104 | 3 | US-09-048-889-5 | Sequence 5, Appli |
| 34 | 6 | 1.7 | 116 | 2 | US-08-960-022-16 | Sequence 16, Appl |
| 35 | 6 | 1.7 | 116 | 6 | 5242821-14 | Patent No. 5242821 |
| 36 | 6 | 1.7 | 127 | 1 | US-08-482-882-45 | Sequence 45, Appl |
| 37 | 6 | 1.7 | 127 | 2 | US-08-483-389-45 | Sequence 45, Appl |
| 38 | 6 | 1.7 | 127 | 2 | US-08-487-113D-45 | Sequence 45, Appl |
| 39 | 6 | 1.7 | 127 | 2 | US-08-473-503-45 | Sequence 45, Appl |
| 40 | 6 | 1.7 | 127 | 2 | US-08-483-932-45 | Sequence 45, Appl |
| 41 | 6 | 1.7 | 127 | 2 | US-08-720-420A-45 | Sequence 45, Appl |
| 42 | 6 | 1.7 | 127 | 3 | US-08-714-017-45 | Sequence 45, Appl |
| 43 | 6 | 1.7 | 127 | 3 | US-08-475-680-45 | Sequence 45, Appl |
| 44 | 6 | 1.7 | 178 | 1 | US-07-929-580B-2 | Sequence 2, Appli |
| 45 | 6 | 1.7 | 179 | 2 | US-08-609-049A-25 | Sequence 25, Appl |

ALIGNMENTS

RESULT 1
US-08-613-220B-4
; Sequence 4, Application US/08613220B
; Patent No. 5958751
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; APPLICANT: Reid, John
; TITLE OF INVENTION: ALPHA-GALACTOSIDASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/613,220B
; FILING DATE: 08-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-68-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 346 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-613-220B-4

Query Match 100.0%; Score 346; DB 2; Length 346;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRALVFHGNLQYAEIPKSEPKVIEKAYIPVETLIKEEPGLNITGYTLKFLPKDILVK 60
Db 1 LRALVFHGNLQYAEIPKSEPKVIEKAYIPVETLIKEEPGLNITGYTLKFLPKDILVK 60
QY 61 GGIASDLIEIIGTSYTAILPLPLSRVEAQVQDRVKEELFEVSPKGFWLPELADPIIPA 120
Db 61 GGIASDLIEIIGTSYTAILPLPLSRVEAQVQDRVKEELFEVSPKGFWLPELADPIIPA 120
QY 121 ILKDNCEYELFADEAMLFSAHLNSAIKPIKPLPHLKAQREKFRYISYLLRELKKAIK 180
Db 121 ILKDNCEYELFADEAMLFSAHLNSAIKPIKPLPHLKAQREKFRYISYLLRELKKAIK 180
QY 181 LVFEKGVTLKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWIEDKNILLYGTDIEFI 240
Db 181 LVFEKGVTLKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWIEDKNILLYGTDIEFI 240
QY 241 GYRDIAGRMSVEGLLEVIDELNSELCPSELKHSRGRELYLRTSSWADKSLRWREDEGNAR 300
Db 241 GYRDIAGRMSVEGLLEVIDELNSELCPSELKHSRGRELYLRTSSWADKSLRWREDEGNAR 300
QY 301 LNMVLMRGELAFLAENSDARGWPLPERRLDAFRAIYNDRWNGEP 346
Db 301 LNMVLMRGELAFLAENSDARGWPLPERRLDAFRAIYNDRWNGEP 346

RESULT 2
US-09-068-195-25
; Sequence 25, Application US/09068195B
; Patent No. 6140078
; GENERAL INFORMATION:
; APPLICANT: Sanders, Jan W.
; APPLICANT: Ledebauer, Adrianus M.
; APPLICANT: Venema, Gerard
; APPLICANT: Kok, Jan
; TITLE OF INVENTION: Salt-Inducible Promoter Derivable from a Lactic Acid
; TITLE OF INVENTION: Bacterium, and Its Use in a Lactic Acid Bacterium for
; TITLE OF INVENTION: Production of Desired Protein
; FILE REFERENCE: Sanders-60113/0252227
; CURRENT APPLICATION NUMBER: US/09/068, 195B
; CURRENT FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: PCT/EP97/04755
; EARLIER FILING DATE: 1997-08-20
; EARLIER APPLICATION NUMBER: EP 97200744/7
; EARLIER FILING DATE: 1997-03-13
; EARLIER APPLICATION NUMBER: EP 96202444/4
; EARLIER FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 466
; TYPE: PRT
; ORGANISM: L. lactis MGL363
US-09-068-195-25

Query Match 2.3%; Score 8; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 295 DEGNARLN 302
Db 52 DEGNARLN 59

RESULT 3
US-08-905-223-439
; Sequence 439, Application US/08905223
; Patent No. 622029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS

; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 439:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -41...-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 7.4
; OTHER INFORMATION: seq ILPLFGCLGVFG/LF
US-08-905-223-439

Query Match 2.0%; Score 7; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 TAILPLL 82
Db 27 TAILPLL 33

RESULT 4
US-09-002-298-3
; Sequence 3, Application US/09002298
; Patent No. 6046001
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Tang, Y. Tom
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN FATTY ACID BETA-OXIDATION ENZYMES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/002,298
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0453 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 301 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: OVRTUT02
; CLONE: 2595635
; US-09-002-298-3
```

Query Match 2.0%; Score 7; DB 3; Length 301;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 SGRELYL 279
| | | | |
Db 263 SGRELYL 269

```
RESULT 5
US-08-913-578-2
; Sequence 2, Application US/08913578
; Patent No. 6218159
; GENERAL INFORMATION:
; APPLICANT: Hodgson, John
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6218159el tRNA synthetase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913,578
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9601069.9
; FILING DATE: 19-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31352
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
```

```
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 484 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-913-578-2
```

Query Match 2.0%; Score 7; DB 4; Length 484;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 LPHLIKA 158
| | | | |
Db 342 LPHLIKA 348

```
RESULT 6
US-08-785-427-2
; Sequence 2, Application us/08785427
; Patent No. 6238900
; GENERAL INFORMATION:
; APPLICANT: Hodgson, John
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6238900el tRNA synthetase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,427
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9601069.9
; FILING DATE: 19-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31352
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 484 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-785-427-2
```

Query Match 2.0%; Score 7; DB 4; Length 484;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 LPHLIKA 158
| | | | |
Db 342 LPHLIKA 348

RESULT 7
US-08-676-279-59
; Sequence 59, Application US/08676279
; Patent No. 5869247
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: MACROPHAGE NUCLEOTIDE SEQUENCE
; NUMBER OF SEQUENCES: 63
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,279
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/00095
; APPLICATION NUMBER: GB 9400929.7
; FILING DATE: 19-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9422021.7
; FILING DATE: 31-OCT-1994
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 549 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-676-279-59

Query Match 2.0%; Score 7; DB 2; Length 549;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 39 PFGLNIT 45
|||||
Db 129 PFGLNIT 135

RESULT 8
US-09-196-293-7
; Sequence 7, Application US/09196293
; Patent No. 6183755
; GENERAL INFORMATION:
; APPLICANT: Fuchs, Renate
; APPLICANT: Motz, Manfred
; APPLICANT: Soutscheck, Erwin
; APPLICANT: Wilske, Bettina
; APPLICANT: Preac-Mursic, Vera
; TITLE OF INVENTION: Active proteins from Borrelia
; TITLE OF INVENTION: burgdorferi
; FILE REFERENCE: 738.001US2
; CURRENT APPLICATION NUMBER: US/09/196,293
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: US 08/209,603
; EARLIER FILING DATE: 1994-03-10
; EARLIER APPLICATION NUMBER: US 07/862,535
; EARLIER FILING DATE: 1992-06-19
; EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
; EARLIER FILING DATE: 1990-12-21
; EARLIER APPLICATION NUMBER: DE P39 42 728.5
; EARLIER FILING DATE: 1989-12-22
; EARLIER APPLICATION NUMBER: DE P40 18 988.0
; EARLIER FILING DATE: 1990-06-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi

US-09-196-293-7
Query Match 1.7%; Score 6; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 250 SVEGLL 255
|||||
Db 5 SVEGLL 10
RESULT 9
US-08-209-603E-7
; Sequence 7, Application US/08209603E
; Patent No. 6248538
; GENERAL INFORMATION:
; APPLICANT: FUCHS, RENATE
; APPLICANT: WILSKE, BETTINA
; APPLICANT: PREAC-MURSIC, VERA
; APPLICANT: MOTZ, MANFRED
; APPLICANT: SOUTSCHECK, ERWIN
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
; TITLE OF INVENTION: FROM BORRELIA BURGDORFERI
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/209,603E
; FILING DATE: 10-MAR-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP90/02282
; FILING DATE: 21-DEC-1990
; APPLICATION NUMBER: US 07/862,535
; FILING DATE: 19-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9217-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PROTEIN
; HYPOTHETICAL: N/A
; ANTI-SENSE: N/A
; FRAGMENT TYPE: N/A
; ORIGINAL SOURCE:
; ORGANISM: B. BURGDORFERI
; IMMEDIATE SOURCE:
; LIBRARY: DSM 5662
; POSITION IN GENOME: N/A
; FEATURE:
; IDENTIFICATION METHOD: amino acid analysis
; PUBLICATION INFORMATION: N/A
US-08-209-603E-7


```

Query Match      1.7%; Score 6; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SVEGLL 255
Db 5 SVEGLL 10

RESULT 10
PCT-US94-10257A-52
; Sequence 52, Application PC/TUS9410257A
; GENERAL INFORMATION:
; APPLICANT: IMMULOGIC PHARMACEUTICAL CORPORATION
; APPLICANT: BRIGITTE DEVAUX
; APPLICANT: JONATHAN B. ROTHBARD
; APPLICANT: DAWN SMILEK
; TITLE OF INVENTION: USES OF MYELIN OLIGODENDROCYTE
; TITLE OF INVENTION: GLYCOPROTEIN AND PEPTIDE PORTIONS THEREOF IN PROTOCOLS
; TITLE OF INVENTION: RELATED TO AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02145
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10257A
; FILING DATE: 1 SEPTEMBER 1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,824
; FILING DATE: 03-SEP-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ANNE I CRAIG
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 071.1 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
PCT-US94-10257A-52

Query Match      1.7%; Score 6; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 EGKVTL 189
Db 5 EGKVTL 10

RESULT 11
US-08-378-761A-48
; Sequence 48, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:

```

```

; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,761A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-378-761A-48

Query Match      1.7%; Score 6; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 RERYIS 168
Db 10 RERYIS 15

RESULT 12
US-08-485-286-48
; Sequence 48, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378761
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-485-286-48

Query Match 1.7%; Score 6; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 RFRYIS 168
|||||
Db 10 RFRYIS 15

RESULT 13
5248606-29
; Patent No. 5248606
; APPLICANT: WALSH, TERENCE A.; HEY, TIMOTHY D.; MORGAN,
; ALICE E.R.
; TITLE OF INVENTION: DNA ENCODING INACTIVE PRECURSOR AND
; ACTIVE FORMS OF MAIZE RIBOSOME INACTIVATIN
; NUMBER OF SEQUENCES: 49
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/535,636
; FILING DATE: 11-JUN-1990
; SEQ ID NO: 29;
; LENGTH: 15
5248606-29

Query Match 1.7%; Score 6; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 RFRYIS 168
|||||
Db 10 RFRYIS 15

RESULT 14
US-08-786-748A-128
; Sequence 128, Application US/08786748A
; Patent No. 5714577
; GENERAL INFORMATION:
; APPLICANT: Ronald, Montelaro C.
; APPLICANT: Tencza, Sarah B.
; APPLICANT: Mietzner, Timothy A.
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA

; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,748A
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/010,634
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rochelle K. Seide
; REGISTRATION NUMBER: 32,300
; REFERENCE/DOCKET NUMBER: AP30421
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2500
; TELEFAX: 212-765-2519
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5714577e
US-08-786-748A-128

Query Match 1.7%; Score 6; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 GLLEVI 258
|||||
Db 3 GLLEVI 8

RESULT 15
US-08-786-748A-132
; Sequence 132, Application US/08786748A
; Patent No. 5714577
; GENERAL INFORMATION:
; APPLICANT: Ronald, Montelaro C.
; APPLICANT: Tencza, Sarah B.
; APPLICANT: Mietzner, Timothy A.
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,748A
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/010,634
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rochelle K. Seide
; REGISTRATION NUMBER: 32,300
; REFERENCE/DOCKET NUMBER: AP30421
; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 212-408-2500
; TELEFAX: 212-765-2519
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5714577e
US-08-786-748A-132

```

```

Query Match      1.7%  Score 6;  DB 1;  Length 18;
Best Local Similarity 100.0%;  Pred. No. 37;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

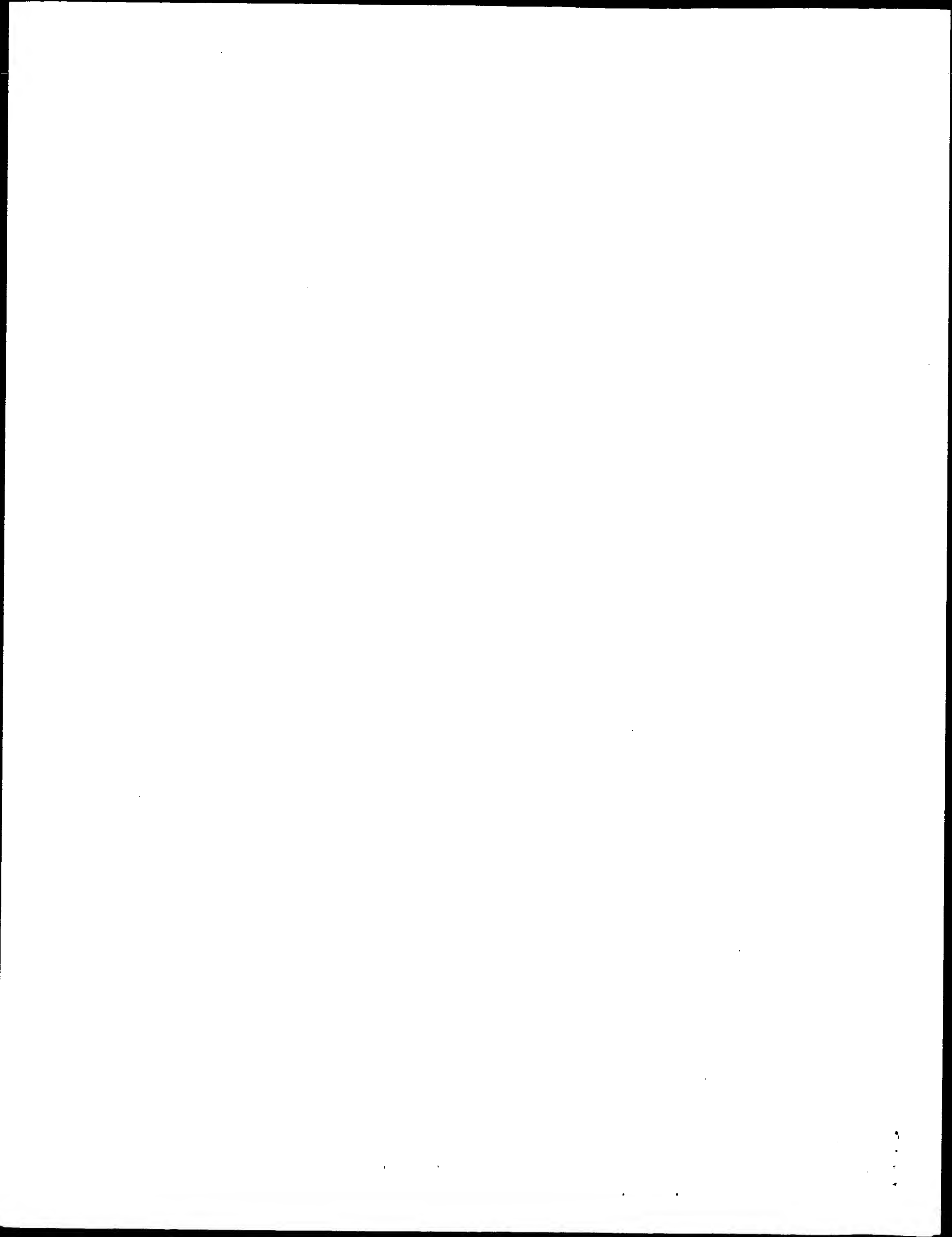
QY  253 GLLEVI 258
    |||||
Db   3 GLLEVI 8

```

```

Search completed: April 4, 2002, 14:46:01
Job time: 170 sec

```



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 14:43:52 ; Search time 21.66 Seconds
(without alignments)
1216.823 Million cell updates/sec

Title: US-09-407-806A-4
Perfect score: 346
Sequence: 1 LRALVFHGNLQYAEIPKSEP.....ERRLDAFRAIYNDWRNGEP 346

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_68:
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 12 | 3.5 | 364 | 2 E71144 | hypothetical prote |
| 2 | 8 | 2.3 | 218 | 2 A64445 | probable phosphoes |
| 3 | 8 | 2.3 | 243 | 2 S75330 | hypothetical prote |
| 4 | 8 | 2.3 | 272 | 2 H75219 | hypothetical prote |
| 5 | 8 | 2.3 | 272 | 2 A71203 | hypothetical prote |
| 6 | 8 | 2.3 | 466 | 2 B86786 | glutamate decarbox |
| 7 | 8 | 2.3 | 492 | 2 T01584 | cellulase (EC 3.2. |
| 8 | 8 | 2.3 | 529 | 2 A47114 | phosphoprotein pho |
| 9 | 8 | 2.3 | 619 | 2 F81532 | conserved hypothet |
| 10 | 8 | 2.3 | 619 | 2 F72006 | CT858 hypothetical |
| 11 | 8 | 2.3 | 619 | 2 E86617 | CT858 hypothetical |
| 12 | 8 | 2.3 | 878 | 2 S47332 | DNA topoisomerase |
| 13 | 7 | 2.0 | 128 | 2 D86809 | ferric uptake regu |
| 14 | 7 | 2.0 | 147 | 1 RPBPF5 | immunity repressor |
| 15 | 7 | 2.0 | 175 | 1 G69518 | isochorismatase (e |
| 16 | 7 | 2.0 | 178 | 1 I40072 | translation initia |
| 17 | 7 | 2.0 | 195 | 2 S34993 | nitrite reductase |
| 18 | 7 | 2.0 | 231 | 2 H64466 | hypothetical prote |
| 19 | 7 | 2.0 | 248 | 2 C83865 | hypothetical prote |
| 20 | 7 | 2.0 | 255 | 2 B70982 | probable nei prote |
| 21 | 7 | 2.0 | 267 | 2 A75131 | hypothetical prote |
| 22 | 7 | 2.0 | 277 | 2 G71456 | hypothetical prote |
| 23 | 7 | 2.0 | 302 | 2 E72402 | phospho-N-acetylmu |
| 24 | 7 | 2.0 | 304 | 1 G69601 | integrase/recombin |
| 25 | 7 | 2.0 | 318 | 2 D82742 | dihydroxydipicolin |
| 26 | 7 | 2.0 | 325 | 2 T17307 | hypothetical prote |
| 27 | 7 | 2.0 | 326 | 2 A83273 | fatty acid biosynt |
| 28 | 7 | 2.0 | 330 | 1 YWBS | tryptophan--trNA l |
| 29 | 7 | 2.0 | 333 | 2 H75637 | probable transposa |

| | | | | | |
|----|---|-----|-----|----------|--------------------|
| 30 | 7 | 2.0 | 352 | 2 T36719 | probable integral |
| 31 | 7 | 2.0 | 358 | 2 JC5964 | apoptosis inhibito |
| 32 | 7 | 2.0 | 365 | 2 S61636 | hypothetical prote |
| 33 | 7 | 2.0 | 374 | 2 T19866 | hypothetical prote |
| 34 | 7 | 2.0 | 380 | 2 T00792 | hypothetical prote |
| 35 | 7 | 2.0 | 389 | 2 B96522 | hypothetical prote |
| 36 | 7 | 2.0 | 392 | 2 D70475 | conserved hypothet |
| 37 | 7 | 2.0 | 397 | 2 T51603 | monooxygenase 1 [i |
| 38 | 7 | 2.0 | 430 | 2 S50981 | probable membrane |
| 39 | 7 | 2.0 | 437 | 2 T28180 | hypothetical prote |
| 40 | 7 | 2.0 | 450 | 2 S73545 | ATP-dependent RNA |
| 41 | 7 | 2.0 | 471 | 2 A64931 | hypothetical prote |
| 42 | 7 | 2.0 | 475 | 2 I76668 | pregnancy-specific |
| 43 | 7 | 2.0 | 496 | 2 B71489 | probable exodeoxyr |
| 44 | 7 | 2.0 | 511 | 1 E70391 | major facilitator |
| 45 | 7 | 2.0 | 539 | 2 T36524 | probable ABC-type |

ALIGNMENTS

RESULT 1

E71144

hypothetical protein PH0368 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 08-Sep-2000

C:Accession: E71144

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Og

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Og

DNA Res. 5, 55-76, 1998

A>Title: Complete sequence and gene organization of the genome of a hyper-thermophil

A:Reference number: A71000; MUID:98344137

A:Accession: E71144

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-364 <KAW>

A:Cross-references: GB:AP000002; NID:g3236129; PIDN:BAA29442.1; PID:dl030385; PID:g3

A:Experimental source: strain OT3

A>Note: this accession replaces an interim accession for a sequence replaced by GenB

C:Genetics:

A:Gene: PH0368

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0368

Query Match 3.5%; Score 12; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 0.00071;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 FHGNLQYAEIPK 17

|||||

Db 6 FHGNLQYAEIPK 17

RESULT 2

A64445

probable phosphoesterase (EC 3.1.1.-) MJ1162 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: A64445

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Bla

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woes

A>Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannas

A:Reference number: A64300; MUID:96337999

A:Accession: A64445

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-218 <BUL>

A:Cross-references: GB:U67558; GB:L77117; NID:g1591786; PIDN:AAB99164.1; PID:g159178

C:Genetics:

A:Map position: REV1103956-1103300

C;Superfamily: probable phosphoesterase MTH1882; phosphoesterase core homology
C;Keywords: hydrolase
F;2-67/Domain: phosphoesterase core homology <PEC>

Query Match 2.3%; Score 8; DB 2; Length 218;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 256 EVIDELNS 263

|||||

Db 70 EVIDELNS 77

RESULT 3

S75330

hypothetical protein sll1570 - *Synechocystis* sp. (strain PCC 6803)

C;Species: *Synechocystis* sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C;Accession: S75330

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
S.

A;Reference number: S74322; MUID:97061201

A;Accession: S75330

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-243 <KAN>

A;Cross-references: EMBL:D90904; GB:AB001339; NID:g1652225; PIDN:BAA17244.1; PID:g165232

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C;Genetics:

A;Start codon: GTG

C;Superfamily: *Synechocystis* hypothetical protein sll1570

Query Match 2.3%; Score 8; DB 2; Length 243;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 PKKVASWI 225

|||||

Db 16 PKKVASWI 23

RESULT 4

H75219

hypothetical protein PAB2165 - *Pyrococcus abyssi* (strain Orsay)

C;Species: *Pyrococcus abyssi*

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 28-Jul-2000

C;Accession: H75219

R;anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A;Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome stru

A;Reference number: A75001

A;Accession: H75219

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-272 <KAN>

A;Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49207.1; PID:g545771

A;Experimental source: strain Orsay

C;Genetics:

A;Gene: PAB2165

C;Superfamily: *Pyrococcus abyssi* hypothetical protein PAB2165

Query Match 2.3%; Score 8; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 254 LLEVIDEL 261

Db 173 LLEVIDEL 180

|||||

RESULT 5

A71203

hypothetical protein PH1894 - *Pyrococcus horikoshii*

C;Species: *Pyrococcus horikoshii*

C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 28-Jul-2000

C;Accession: A71203

R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A;Reference number: A71000; MUID:98344137

A;Accession: A71203

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-272 <KAW>

A;Cross-references: GB:AP000007; NID:g23236134; PIDN:BAA31016.1; PID:g3258333

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBa

C;Genetics:

A;Gene: PH1894

C;Superfamily: *Pyrococcus abyssi* hypothetical protein PAB2165

Query Match 2.3%; Score 8; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 254 LLEVIDEL 261

|||||

Db 173 LLEVIDEL 180

RESULT 6

B86786

glutamate decarboxylase (EC 4.1.1.15) [imported] - *Lactococcus lactis* subsp. *lactis* (

C;Species: *Lactococcus lactis* subsp. *lactis*

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001

C;Accession: B86786

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Eh
Genome Res. in press, 2001

A;Title: The complete genome sequence of the lactic acid bacterium.

A;Reference number: A86625

A;Accession: B86786

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-466 <STO>

A;Cross-references: GB:AE005176; NID:g12724267; PIDN:AAK05388.1; GSPDB:GN00146

A;Experimental source: strain ILL403

C;Genetics:

A;Gene: gadB

C;Superfamily: *Escherichia coli* glutamate decarboxylase

C;Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 2.3%; Score 8; DB 2; Length 466;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 DEGNARLN 302

|||||

Db 52 DEGNARLN 59

RESULT 7

T01584

cellulase (EC 3.2.1.4) F16B22.6 - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 16-Feb-2001

C;Accession: T01584; B84880

R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, July 1998
A;Description: Arabidopsis thaliana chromosome II BAC F16B22 genomic sequence.
A;Reference number: Z14284
A;Accession: T01584
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-492 <ROU>
A;Cross-references: EMBL:AC003672; NID:g3341671; PID:g3341677
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487
A;Accession: B84880
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-492 <STO>
A;Cross-references: GB:AE002093; NID:g3341677; PIDN:AAC27459.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g44570; F16B22.6
A;Map position: 2
A;Introns: 74/3; 292/3; 387/3
C;Superfamily: Arabidopsis membrane-anchored cellulase KOR
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 2.3%; Score 8; DB 2; Length 492;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 KLVFEGKV 187
|
Db 292 KLVFEGKV 299
|

RESULT 8
A47114
phosphoprotein phosphatase (EC 3.1.3.16) 2A regulatory chain PR72 - human
C;Species: Homo sapiens (man)
C;Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 08-Oct-1999
C;Accession: A47114
R;Hendrix, P.; Mayer-Jaekel, R.E.; Cron, P.; Goris, J.; Hofsteenge, J.; Merlevede, W.; H. J. Biol. Chem. 268, 15267-15276, 1993
A;Title: Structure and expression of a 72-kDa regulatory subunit of protein phosphatase
A;Reference number: A47114; MUID:93315512
A;Accession: A47114
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-529 <HEN>
A;Cross-references: GB:L12146; NID:g190221; PIDN:AAB02614.1; PID:g190222
C;Keywords: alternative splicing; phosphoric monoester hydrolase

Query Match 2.3%; Score 8; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 RGEAFLA 315
|
Db 15 RGEAFLA 22
|

RESULT 9
F81532
conserved hypothetical protein CP0837 [imported] - Chlamydomophila pneumoniae (strain AR39
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: F81532
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255
A;Accession: F81532
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-619 <REA>
A;Cross-references: GB:AE002243; GB:AE002161; NID:g7189750; PIDN:AAF38630.1; PID:g718
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0837

Query Match 2.3%; Score 8; DB 2; Length 619;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 PAALKDNG 126
|
Db 498 PAALKDNG 505
|

RESULT 10
F72006
CT858 hypothetical protein - Chlamydomophila pneumoniae (strain CWL029)
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C;Accession: F72006
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606
A;Accession: F72006
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-619 <ARN>
A;Cross-references: GB:AE001682; GB:AE001363; NID:g4377344; PIDN:AAD19153.1; PID:g437
A;Experimental source: strain CWL029
C;Genetics:
A;Gene: CPn1016

Query Match 2.3%; Score 8; DB 2; Length 619;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 PAALKDNG 126
|
Db 498 PAALKDNG 505
|

RESULT 11
E86617
CT858 hypothetical protein [imported] - Chlamydomophila pneumoniae (strain J138)
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C;Accession: E86617
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349
A;Accession: E86617
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-619 <STO>
A;Cross-references: GB:BA000008; NID:g8979389; PIDN:BAA99223.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: CPj1016

Query Match 2.3%; Score 8; DB 2; Length 619;
Best Local Similarity 100.0%; Pred. No. 12;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 PAIKDNG 126
|||||||

Db 498 PAIKDNG 505

RESULT 12

S47332

DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) - Erwinia carotovora

C;Species: Erwinia carotovora

C;Date: 23-Nov-1994 #sequence_revision 03-Aug-1995 #text_change 16-Jul-1999

C;Accession: S47332

R;Rosanas, A.

submitted to the EMBL Data Library, August 1994

A;Reference number: S47332

A;Accession: S47332

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-878 <ROS>

A;Cross-references: EMBL:X80798; NID:g525201; PIDN:CAA56777.1; PID:g525202

C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase

C;Keywords: ATP; isomerase

F;1-238/Domain: phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology <T4T>

Query Match

Best Local Similarity 2.3%; Score 8; DB 2; Length 878;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DEAMLFSA 140

|||||||

Db 691 DEAMLFSA 698

RESULT 13

D86809

ferric uptake regulator [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C;Species: Lactococcus lactis subsp. lactis

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001

C;Accession: D86809

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlich

Genome Res. in press, 2001

A;Title: The complete genome sequence of the lactic acid bacterium.

A;Reference number: A86625

A;Accession: D86809

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-128 <STO>

A;Cross-references: GB:AE005176; NID:g12724470; PIDN:AAK05574.1; GSPDB:GN00146

A;Experimental source: strain IL1403

C;Genetics:

A;Gene: fur

Query Match

Best Local Similarity 2.0%; Score 7; DB 2; Length 128;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 GLNITGY 47

|||||||

Db 113 GLNITGY 119

RESULT 14

RBPBF5

immunity repressor protein - Bacillus phage phi-105

C;Species: Bacillus phage phi-105

A;Note: host Bacillus subtilis

C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jun-2000

C;Accession: A93579; A91535; S02459; T13543; A24339; A24521

R;Dhaese, P.; Seurinck, J.; De Smet, B.; Van Montagu, M.

Nucleic Acids Res. 13, 5441-5455, 1985

A;Title: Nucleotide sequence and mutational analysis of an immunity repressor gene fr

A;Reference number: A93579; MUID:85297750

A;Accession: A93579

A;Molecule type: DNA

A;Residues: 1-147 <DHA>

A;Cross-references: GB:X02799; NID:g15455; PIDN:CAA26567.1; PID:g579178

R;Cully, D.F.; Garro, A.J.

Gene 38, 153-164, 1985

A;Title: Nucleotide sequence of the immunity region of Bacillus subtilis bacteriophag

A;Reference number: A91535; MUID:86056972

A;Accession: A91535

A;Molecule type: DNA

A;Residues: 1-147 <CUL>

R;van Kaer, L.; Gansemans, Y.; van Montagu, M.; Dhaese, P.

EMBO J. 7, 859-866, 1988

A;Title: Interaction of the Bacillus subtilis phage phi105 repressor with operator DN

A;Reference number: S02459; MUID:88283656

A;Accession: S02459

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 4-147 <VAN>

A;Note: part of this sequence, including the amino end of the mature protein, was con

R;Kobayashi, K.; Okamura, K.; Inoue, T.; Sato, T.; Kobayashi, Y.

submitted to the EMBL Data Library, July 1998

A;Description: Complete nucleotide sequence of Bacillus subtilis phage phi-105.

A;Reference number: Z17688

A;Accession: T13543

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 4-147 <KOB>

A;Cross-references: EMBL:AB016282; PIDN:BAA36660.1

C;Genetics:

A;Start codon: GTG

C;Superfamily: phage phi-105 immunity repressor protein

C;Keywords: DNA binding; early protein; repressor; transcription regulation

Query Match

Best Local Similarity 2.0%; Score 7; DB 1; Length 147;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 ETLIKEE 38

|||||||

Db 68 ETLIKEE 74

RESULT 15

G69518

isochorismatase (entB) homolog - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C;Accession: G69518

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod

Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch

A;Reference number: A69250; MUID:98049343

A;Accession: G69518

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-175 <KLE>

A;Cross-references: GB:AE000956; GB:AE000782; NID:g2689279; PIDN:AAB89104.1; PID:g264

C;Superfamily: hypothetical protein b1011

Query Match

Best Local Similarity 2.0%; Score 7; DB 1; Length 175;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 256 EVIDELN 262

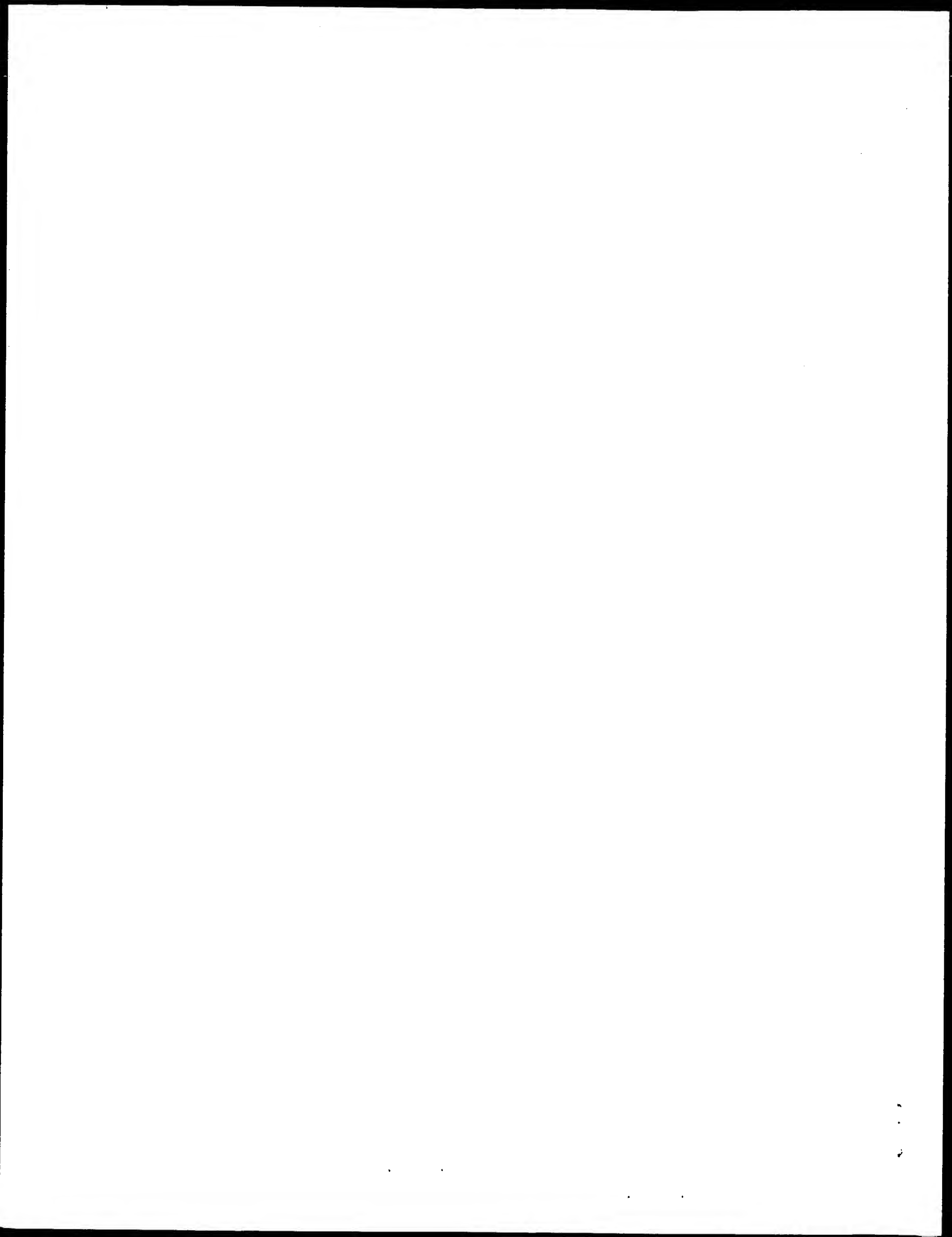
Thu Apr 4 16:34:54 2002

us-09-407-806a-4.rpr

Page 5

Db |||||||
 74 EVIDELN 80

Search completed: April 4, 2002, 14:46:37
Job time: 165 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 14:46:02 ; Search time 12.23 Seconds
(without alignments)
1037.288 Million cell updates/sec

Title: US-09-407-806A-4
Perfect score: 346
Sequence: 1 LRALVFHGNLQYAEIPKSEP.....ERRLDAFRAIYNDRNGNEP 346

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 100059 seqs, 36664827 residues
Word size : 0
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------|--------------------|
| 1 | 8 | 2.3 | 619 | 1 YA16_CHLPN | Q9Z6P3 chlamydia p |
| 2 | 8 | 2.3 | 878 | 1 GYRA_ERWCA | P41513 erwina car |
| 3 | 7 | 2.0 | 144 | 1 RPC_BPPHI | P06153 bacterioph |
| 4 | 7 | 2.0 | 151 | 1 SYB2_RHIME | Q9Z3Q1 rhizobium m |
| 5 | 7 | 2.0 | 178 | 1 IF3_BUCAP | P46243 buchnera ap |
| 6 | 7 | 2.0 | 231 | 1 YD37_METJA | Q58733 methanococc |
| 7 | 7 | 2.0 | 237 | 1 Y4MC_RHISN | P55562 rhizobium s |
| 8 | 7 | 2.0 | 255 | 1 END8_MYCTU | P96902 mycobacteri |
| 9 | 7 | 2.0 | 302 | 1 DAPA_XYLEFA | Q9PER5 xylella fas |
| 10 | 7 | 2.0 | 302 | 1 MRAY_THEMEA | Q9WY77 thermotoga |
| 11 | 7 | 2.0 | 304 | 1 CODV_BACSU | P39776 bacillus su |
| 12 | 7 | 2.0 | 321 | 1 LPXK_RICRI | P58187 rickettsia |
| 13 | 7 | 2.0 | 325 | 1 LPXK_RICMO | P58186 rickettsia |
| 14 | 7 | 2.0 | 330 | 1 SYW_BACSU | P21656 bacillus su |
| 15 | 7 | 2.0 | 354 | 1 GBA3_USTHO | O14438 ustilago ho |
| 16 | 7 | 2.0 | 358 | 1 PIAP_PIG | O62640 sus scrofa |
| 17 | 7 | 2.0 | 450 | 1 Y425_MYCPN | P75172 mycoplasma |
| 18 | 7 | 2.0 | 549 | 1 SMF2_YEAST | P38778 saccharomyc |
| 19 | 7 | 2.0 | 601 | 1 Y248_CHLMU | Q9P160 chlamydia m |
| 20 | 7 | 2.0 | 632 | 1 ARPB_ECOLI | P76205 escherichia |
| 21 | 7 | 2.0 | 644 | 1 YED0_YEAST | P39987 saccharomyc |
| 22 | 7 | 2.0 | 666 | 1 YB9G_YEAST | P38339 saccharomyc |
| 23 | 7 | 2.0 | 709 | 1 UVRB_MICLU | P10125 micrococcus |
| 24 | 7 | 2.0 | 815 | 1 RPGR_HUMAN | Q92834 homo sapien |
| 25 | 7 | 2.0 | 1237 | 1 POL4_DROME | P10394 drosophila |
| 26 | 7 | 2.0 | 1320 | 1 PUR4_NEIMB | Q9JXK5 neisseria m |
| 27 | 7 | 2.0 | 2499 | 1 MPRI_BOVIN | P08169 bos taurus |
| 28 | 7 | 2.0 | 4684 | 1 PLE1_HUMAN | Q15149 homo sapien |
| 29 | 7 | 2.0 | 4687 | 1 PLE1_RAT | P30427 rattus norv |
| 30 | 6 | 1.7 | 24 | 1 ALKB_SALTY | P37462 salmonella |
| 31 | 6 | 1.7 | 63 | 1 ANP1_AUSBR | P12100 austrolcyc |
| 32 | 6 | 1.7 | 63 | 1 ANP2_AUSBR | P12101 austrolcyc |
| 33 | 6 | 1.7 | 63 | 1 RL29_VIBCH | Q9KNZ2 vibrio chol |

| | | | | | |
|----|---|-----|-----|--------------|--------------------|
| 34 | 6 | 1.7 | 70 | 1 RL29_HALMA | PL10971 haloarcula |
| 35 | 6 | 1.7 | 94 | 1 CH10_THEBR | Q60023 thermoanaer |
| 36 | 6 | 1.7 | 105 | 1 YNC8_YEAST | P53967 saccharomyc |
| 37 | 6 | 1.7 | 107 | 1 QACG_STASP | O87866 staphylococ |
| 38 | 6 | 1.7 | 112 | 1 GLNB_PORPU | P51254 porphyra pu |
| 39 | 6 | 1.7 | 118 | 1 Y493_HAEIN | O05023 haemophilus |
| 40 | 6 | 1.7 | 125 | 1 RS6_PASMU | Q9CLN8 pasteurella |
| 41 | 6 | 1.7 | 128 | 1 C562_ECOLI | P00192 escherichia |
| 42 | 6 | 1.7 | 128 | 1 RS6_THEMEA | Q9WZ72 thermotoga |
| 43 | 6 | 1.7 | 129 | 1 EXBD_HAEDU | O51809 haemophilus |
| 44 | 6 | 1.7 | 133 | 1 YBB2_YEAST | P38209 saccharomyc |
| 45 | 6 | 1.7 | 134 | 1 ANP3_RHIDE | P35753 rhigophila |

ALIGNMENTS

RESULT 1
YA16_CHLPN STANDARD; PRT; 619 AA.
AC Q9Z6P3; Q9JS57;
DT 30-MAY-2000 (Rel. 39, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN CPN1016/CP0837/CPJ1016.
GN CPN1016 OR CP0837 OR CPJ1016.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN1016/CT858/TC0248
FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AE001682; AAD19153.1; -
EMBL; AE002243; AAF38630.1; -
EMBL; AP002548; BAA99223.1; -
TIGR; CP0837; -

DR InterPro; IPR000178; IF2.
DR InterPro; IPR003581; TSPC.
DR ProDom; PD186100; IF2; 1.
DR SMART; SM00245; TSPC; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 619 619 L -> S (IN REF. 1).
SQ SEQUENCE 619 AA; 69178 MW; 25C3B607B5936FDB CRC64;

Query Match 2.3%; Score 8; DB 1; Length 619;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 PAILKDNG 126
|||||
Db 498 PAILKDNG 505

RESULT 2
GYRA_ERWCA STANDARD; PRT; 878 AA.
AC P41513;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE DNA GYRASE SUBUNIT A (EC 5.99.1.3).
GN GYRA.
OS Erwinia carotovora.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pectobacterium.
OX NCBI_TaxID=554;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95369719; PubMed=7642123;
RA Rosanas A., Barbe J., Gibert I.;
RT "Cloning and sequencing of the gyra gene from the plant pathogen
RT Erwinia carotovora.";
RL Gene 161:11-14(1995).

CC -!- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
CC OF DOUBLE-STRANDED DNA.
CC -!- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
CC ENZYME FORMS AN A2B2 TETRAMER.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X80798; CAA56777.1; -.
DR InterPro; IPR002205; DNA_topoisolv.
DR Pfam; PF00521; DNA_topoisolv; 1.
DR SMART; SM00434; TOP4c; 1.
KW Topoisomerase; Isomerase; DNA-binding.
FT ACT_SITE 122 122 DNA CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 878 AA; 97588 MW; 6F7D181BADA617E3 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 878;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DEAMLFSA 140
|||||
Db 691 DEAMLFSA 698

RESULT 3
RPC_BPPH1

ID RPC_BPPH1 STANDARD; PRT; 144 AA.
AC P06153; P15239;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE IMMUNITY REPRESSOR PROTEIN.
OS Bacteriophage phi-105.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OC Lambda phage group.
OX NCBI_TaxID=10717;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=88283656; PubMed=3135184;
RA van Kaer L., Gansemans Y., van Montagu M., Dhaese P.;
RT "Interaction of the Bacillus subtilis phage phi 105 repressor DNA: a
RT genetic analysis.";
RL EMBO J. 7:859-866(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85297750; PubMed=2993999;
RA Dhaese P., Seurinck J., de Smet B., van Montagu M.;
RT "Nucleotide sequence and mutational analysis of an immunity repressor
RT gene from Bacillus subtilis temperate phage phi 105.";
RL Nucleic Acids Res. 13:5441-5455(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86056972; PubMed=3934047;
RA Cully D.F., Garro A.J.;
RT "Nucleotide sequence of the immunity region of Bacillus subtilis
RT bacteriophage phi 105: identification of the repressor gene and its
RT mRNA and protein products.";
RL Gene 38:153-164(1985).

CC -!- SIMILARITY: BELONGS TO THE PBSX(XRE) FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X02799; CAA26567.1; ALT_INIT.
DR EMBL; M11920; AAA88396.1; -.
DR PIR; A24339; RPBPF5.
DR PIR; S02459; S02459.
DR InterPro; IPR001387; HTH_3.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Transcription regulation; Repressor; DNA-binding; Early protein.
FT DNA_BIND 18 37 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 144 AA; 16520 MW; EEED6EB3E4B34A1E CRC64;

Query Match 2.0%; Score 7; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 ETLIKEE 38
|||||
Db 65 ETLIKEE 71

RESULT 4
SYB2_RHIME STANDARD; PRT; 151 AA.
ID SYB2_RHIME
AC Q923Q1;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE TRANSCRIPTIONAL REGULATOR SYRB2.
GN SYRB2.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CXMI-105;
RX MEDLINE=99413305; PubMed=10485295;
RA Sharypova L.A., Yurgel S.N., Keller M., Simarov B.V., Puehler A.,
RA Becker A.;
RT "The eff-482 locus of Sinorhizobium meliloti CXMI-105 that influences
RT symbiotic effectiveness consists of three genes encoding an
RT endoglycanase, a transcriptional regulator and an adenylate cyclase.";
RL Mol. Gen. Genet. 261:1032-1044(1999).
CC -!- FUNCTION: SEEMS TO AFFECT THE TRANSCRIPTION OF CYA3. MAY BE
CC -!- NEGATIVELY AUTOREGULATED.
CC -!- SIMILARITY: BELONGS TO THE SYRB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ225896; CAB38102.1; -.
KW Nodulation; Transcription regulation; Repressor; Plasmid.
SQ SEQUENCE 151 AA; 16694 MW; 08A85DC217812BFC CRC64;

Query Match 2.0%; Score 7; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 EGKVTLK 190
| | | | |
pB 73 EGKVTLK 79

RESULT 5
IF3_BUCAP STANDARD; PRT; 178 AA.
AC P46243;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TRANSLATION INITIATION FACTOR IF-3.
GN INFC.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95245268; PubMed=7766160;
RA Kolibachuk D., Rouhakhsh D., Baumann P.;
RT "Aromatic amino acid biosynthesis in Buchnera aphidicola (endosymbiont
RT of aphids): cloning and sequencing of a DNA fragment containing
RT aroH-thrS-infC-rpmI-rplT";
RL Curr. Microbiol. 30:313-316(1995).
CC -!- FUNCTION: IF-3 BINDS TO THE 30S RIBOSOMAL SUBUNIT AND SHIFTS THE
CC EQUILIBRIUM BETWEEN 70S RIBOSOMES AND THEIR 50S AND 30S SUBUNITS IN
CC FAVOR OF THE FREE SUBUNITS, THUS ENHANCING THE AVAILABILITY OF 30S
CC SUBUNITS ON WHICH PROTEIN SYNTHESIS INITIATION BEGINS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE IF-3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U11066; AAC43606.1; -.
DR HSSP; P02999; 2IFE.
DR InterPro; IPR001288; IF3.
DR Pfam; PF00707; IF3; 1.
DR ProDom; PD002880; IF3; 1.
DR PROSITE; PS00938; IF3; 1.
KW Initiation factor; protein biosynthesis.
FT SITE 106 106 IMPORTANT FOR 30S BINDING
FT SITE 109 109 (BY SIMILARITY).
FT SITE 109 109 IMPORTANT FOR 30S BINDING
FT SITE 109 109 (BY SIMILARITY).
SQ SEQUENCE 178 AA; 20601 MW; 9365C9D6512B1A3E CRC64;

Query Match 2.0%; Score 7; DB 1; Length 178;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 LRELKKA 178
| | | | |
pB 39 LRELKKA 45

RESULT 6
YD37_METJA STANDARD; PRT; 231 AA.
ID YD37_METJA
AC Q58733;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN MJ1337.
GN MJ1337.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67574; AAB99358.1; -.
DR TIGR; MJ1337; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 231 AA; 25674 MW; 203926DBC472FAE CRC64;

Query Match 2.0%; Score 7; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 309 GELAFILA 315
Db 94 GELAFILA 100

RESULT 7
Y4MC_RHISN STANDARD; PRT; 237 AA.
AC P55562;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL PROTEIN Y4MC PRECURSOR.
GN Y4MC.
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -!- SIMILARITY: NONE OBVIOUS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AE000085; AAB91766.1; --
KW Hypothetical protein; Plasmid; Signal.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 237 HYPOTHETICAL PROTEIN Y4MC.
SQ SEQUENCE 237 AA; 26205 MW; F5AA67D9EB7BFB5B CRC64;

Query Match 2.0%; Score 7; DB 1; Length 237;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 PLSRVEA 89
Db 41 PLSRVEA 47

RESULT 8
END8_MYCTU STANDARD; PRT; 255 AA.
AC P96902;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE ENDONUCLEASE VIII (EC 3.2.-.-).
GN NEI OR RV3297 OR MT3396 OR MTCY71.37.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekaita F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: DNA N-GLYCOSYLASE WITH AN AP LYASE ACTIVITY. REQUIRED
CC FOR THE REPAIR OF OXIDATIVE DNA DAMAGE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE FPG FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; Z92771; CAB07061.1; --
DR EMBL; AE007148; AAK47739.1; --
DR TIGR; MT3396;
DR TubercuList; RV3297;
DR InterPro; IPR000191; Fapy_DNA_glyco.
DR InterPro; IPR000214; Fapy_DNAglyco_zn.
DR Pfam; PF01149; Fapy_DNA_glyco; 1.
DR ProDom; PD003680; Fapy_DNA_glyco; 1.
DR PROSITE; PS01242; FPG; 1.
KW DNA repair; Hydrolase; Glycosidase; Endonuclease; Zinc;
KW Complete proteome.

FT ZN_FING 230 253 POTENTIAL.
SQ SEQUENCE 255 AA; 28525 MW; 2C97B16CD7CD4002 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 GLEVID 259
Db 106 GLEVID 112

RESULT 9
DAPA_XYLFA STANDARD; PRT; 302 AA.
AC Q9PER5;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DIHYDRODIPICOLINATE SYNTHASE (EC 4.2.1.52) (DHDPS).
GN DAPA OR XF0963.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159(2000).
CC -!- CATALYTIC ACTIVITY: L-ASPARTATE 4-SEMIALDEHYDE + PYRUVATE =
CC DIHYDRODIPICOLINATE + 2 H(2)O.
CC -!- PATHWAY: BIOSYNTHESIS OF DIAMINOPIMELATE AND LYSINE FROM ASPARTATE
CC SEMIALDEHYDE; FIRST STEP.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE DHDPS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AE003934; AAF83773.1; ALT_INIT.
CC InterPro; IPR002220; DHDPS.
CC Pfam; PF00701; DHDPS; 1.
CC PRINTS; PR00146; DHPICNTTHASE.
CC PRODOM; PD001859; DHDPS; 1.
CC PROSITE; PS00665; DHDPS_1; 1.
CC PROSITE; PS00666; DHDPS_2; 1.
CC Lyase; Diaminopimelate biosynthesis; Lysine biosynthesis;
KW Complete proteome.
FT ACT_SITE 162 162 BY SIMILARITY.
FT SEQUENCE 302 AA; 31525 MW; 2298992B64D8EC70 CRC64;
SQ

Query Match 2.0%; Score 7; DB 1; Length 302;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 LPLPLS 85
Db 268 LPLPLS 274
|||||

RESULT 10
MRAY_THEME STANDARD; PRT; 302 AA.
ID MRAY_THEME
AC Q9WY77;
DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PHOSPHO-N-ACETYLMURAMOYL-PENTAPEPTIDE-TRANSFERASE (EC 2.7.8.13) (UDP-
DE MURNA-PENTAPEPTIDE PHOSPHOTRANSFERASE).
GN MRAY OR TM0235.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima";
RL Nature 399:323-329(1999).
CC -!- FUNCTION: FIRST STEP OF THE LIPID CYCLE REACTIONS IN THE
CC BIOSYNTHESIS OF THE CELL WALL PEPTIDOGLYCAN (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: UDP-N-ACETYLMURAMOYL-L-ALANYL-D-GLUTAMYL-L-
CC LYSYL-D-ALANYL-D-ALANINE + UNDECAPRENYL PHOSPHATE = UMP +
CC N-ACETYLMURAMOYL-L-ALANYL-D-GLUTAMYL-L-LYSYL-D-ALANYL-D-ALANINE-
CC DIPHOUPHOUNDECAPRENOL.
CC -!- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 4. MRAY
CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AE001707; AAD35326.1; -
CC TIGR; TM0235; -
CC InterPro; IPR000715; Glycos_transf_4.
CC Pfam; PF00953; Glycos_transf_4; 1.
CC PROSITE; PS01347; MRAY_1; 1.
CC PROSITE; PS01348; MRAY_2; 1.
CC Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
KW Complete proteome.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 67 87 POTENTIAL.
FT TRANSMEM 95 115 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT TRANSMEM 204 224 POTENTIAL.
FT TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 281 301 POTENTIAL.
FT SEQUENCE 302 AA; 33814 MW; BB8PF74FEA9205CB CRC64;
SQ

Query Match 2.0%; Score 7; DB 1; Length 302;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 IPVIETL 34
Db 237 IPVIETL 243
|||||

RESULT 11
CODV_BACSU STANDARD; PRT; 304 AA.
ID CODV_BACSU


```
AC P39776;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE INTEGRASE/RECOMBINASE CODV.
GN CODV.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=95302982; PubMed=7783641;
RA Slack F.J., Serror P., Joyce E., Sonenshein A.L.;
RT "A gene required for nutritional repression of the Bacillus subtilis
RT dipeptide permease operon.";
RL Mol. Microbiol. 15:689-702(1995).
RN [2]
RP SEQUENCE OF 1-85 FROM N.A.
RC STRAIN=168;
RA Foulger D., Errington J.;
RT "Cloning and sequencing 7.5Kbp of DNA from Bacillus subtilis upstream
RT of the codV gene.";
RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE "PHAGE" INTEGRASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U13634; AAB03369.1; -.
DR EMBL; AJ000975; CAA04424.1; -.
DR EMBL; Z99112; CAB13487.1; -.
DR HSSP; P21891; 1A0P.
DR Subtilist; BG10965; codV.
DR InterPro; IPR002104; Phage_integrase.
DR Pfam; PF00589; Phage_integrase; 1.
KW DNA recombination; DNA integration; Complete proteome.
FT ACT_SITE 281 281 TRANSIENT COVALENT LINKAGE TO DNA DURING
FT STRAND CLEAVAGE AND REJOINING (BY
FT SIMILARITY).
FT SEQUENCE 304 AA; 35342 MW; B0778DBDA065A053 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 304;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 EELFEVS 104
Db 120 EELFEVS 126
|||||

RESULT 12
LPXK_RICRI
ID LPXK_RICRI STANDARD; PRT; 321 AA.
AC P58187;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TETRAACYCLIDISACCHARIDE 4'-KINASE (EC 2.7.1.130) (LIPID A 4'-KINASE).
GN LPXK.
OS Rickettsia rickettsii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=783;
RN [1]
RP SEQUENCE FROM N.A.

This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
RC STRAIN=84-21C;
RX MEDLINE=21219194; PubMed=11319266;
RA Andersson J.O., Andersson S.G.E.;
RT "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes.";
RL Mol. Biol. Evol. 18:829-839(2001).
CC -!- FUNCTION: TRANSFERS THE GAMMA-PHOSPHATE OF ATP TO THE 4'-POSITION
CC OF A TETRAACYCLIDISACCHARIDE 1-PHOSPHATE INTERMEDIATE (TERMED DS-1-
CC P) TO FORM TETRAACYCLIDISACCHARIDE 1,4'-BIS-PHOSPHATE (LIPID IVA)
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + 2,3-BIS(3-HYDROXYTETRADECANOYL)-D-
CC GLUCOSAMINYL-(BETA-D-1,6)-2,3-BIS(3-HYDROXYTETRADECANOYL)-D-
CC GLUCOSAMINYL BETA-PHOSPHATE = ADP + 2,3,2',3'-TETRAKIS(3-
CC HYDROXYTETRADECANOYL)-D-GLUCOSAMINYL-1,6-BETA-D-GLUCOSAMINE 1,4'-
CC BISPHOSPHATE.
CC -!- PATHWAY: LIPID A BIOSYNTHESIS; SIXTH STEP.
CC -!- SIMILARITY: BELONGS TO THE LPXK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ293329; CAC33716.1; -.
KW Transferase; Kinase; Lipid A biosynthesis; Lipid synthesis;
KW ATP-binding.
FT NP_BIND 54 61 ATP (POTENTIAL).
FT SEQUENCE 321 AA; 36083 MW; 45E6CEB6CC50C0AE CRC64;

Query Match 2.0%; Score 7; DB 1; Length 321;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 GLNITGY 47
Db 247 GLNITGY 253
|||||

RESULT 13
LPXK_RICMO
ID LPXK_RICMO STANDARD; PRT; 325 AA.
AC P58186;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TETRAACYCLIDISACCHARIDE 4'-KINASE (EC 2.7.1.130) (LIPID A 4'-KINASE).
GN LPXK.
OS Rickettsia montana.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=33991;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21219194; PubMed=11319266;
RA Andersson J.O., Andersson S.G.E.;
RT "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes.";
RL Mol. Biol. Evol. 18:829-839(2001).
CC -!- FUNCTION: TRANSFERS THE GAMMA-PHOSPHATE OF ATP TO THE 4'-POSITION
CC OF A TETRAACYCLIDISACCHARIDE 1-PHOSPHATE INTERMEDIATE (TERMED DS-1-
CC P) TO FORM TETRAACYCLIDISACCHARIDE 1,4'-BIS-PHOSPHATE (LIPID IVA)
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + 2,3-BIS(3-HYDROXYTETRADECANOYL)-D-
CC GLUCOSAMINYL-(BETA-D-1,6)-2,3-BIS(3-HYDROXYTETRADECANOYL)-D-
CC GLUCOSAMINYL BETA-PHOSPHATE = ADP + 2,3,2',3'-TETRAKIS(3-
CC HYDROXYTETRADECANOYL)-D-GLUCOSAMINYL-1,6-BETA-D-GLUCOSAMINE 1,4'-
CC BISPHOSPHATE.
CC -!- PATHWAY: LIPID A BIOSYNTHESIS; SIXTH STEP.
CC -!- SIMILARITY: BELONGS TO THE LPXK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ293330; CAC33651.1; -
KW Transferase; Kinase; Lipid A biosynthesis; Lipid synthesis;
KW ATP-binding.
FT NP_BIND 54 61 ATP (POTENTIAL).
SQ SEQUENCE 325 AA; 36481 MW; F59126EC871BD9FB CRC64;

Query Match 2.0%; Score 7; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 GLNITGY 47
|||||
Db 247 GLNITGY 253

RESULT 14
SYW_BACSU STANDARD; PRT; 330 AA.
AC P21656;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TRYPTOPHANYL-TRNA SYNTHETASE (EC 6.1.1.2) (TRYPTOPHAN--TRNA LIGASE)
DE (TRPRS).
GN TRPS.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89211991; PubMed=3149612;
RA Chow K.C., Wong T.F.;
RT "Cloning and nucleotide sequence of the structural gene coding for
RT Bacillus subtilis tryptophanyl-trna synthetase.";
RL Gene 73:537-574(1988).
RN [2]
RP SEQUENCE OF 1-32.
RX MEDLINE=89174673; PubMed=2494170;
RA Xu Z.J., Love M.L., Ma L.Y.Y., Blum M., Bronskill P.M., Bernstein J.,
RA Grey A.A., Hofmann T., Camerman N., Wong J.T.F.;
RT "Tryptophanyl-trna synthetase from Bacillus subtilis.
RT Characterization and role of hydrophobicity in substrate
RT recognition.";
RL J. Biol. Chem. 264:4304-4311(1989).
CC -!- CATALYTIC ACTIVITY: ATP + L-TRYPTOPHAN + TRNA(TRP) = AMP +
CC PYROPHOSPHATE + L-TRYPTOPHANYL-TRNA(TRP).
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M24068; AAA22874.1; -
DR EMBL; Z99110; CAB12999.1; -
DR PIR; JT0481; YWBS.
DR Subtilist; BG10799; trpS.
DR InterPro; IPR002305; trna-synt_lb.
DR InterPro; IPR001412; trna-synt_I.

DR InterPro; IPR002306; trna-synt_trp.
DR Pfam; PF00579; trna-synt_lb; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 11 19 "HIGH" REGION.
FT SITE 193 197 "KMSKS" REGION.
FT BINDING 196 196 ATP (BY SIMILARITY).
SQ SEQUENCE 330 AA; 37197 MW; 31927E2BC6320544 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 ILLYGTD 236
|||||
Db 134 ILLYGTD 140

RESULT 15
GBA3_USTHO STANDARD; PRT; 354 AA.
ID GBA3_USTHO
AC O14438;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GUANINE NUCLEOTIDE-BINDING PROTEIN ALPHA-3 SUBUNIT.
GN FIL1.
OS Ustilago hordei (Smut fungus).
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=120017;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UH100;
RX MEDLINE=98055158; PubMed=9393440;
RA Lichter A., Mills D.;
RT "Fill, a G-protein alpha-subunit that acts upstream of CAMP and is
RT essential for dimorphic switching in haploid cells of Ustilago
RT hordei.";
RL Mol. Gen. Genet. 256:426-435(1997).
CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS. THIS SUBUNIT IS INVOLVED IN CAMP REGULATION AND
CC MORPHOGENESIS. IT IS ESSENTIAL FOR DIMORPHIC SWITCHING IN HAPLOID
CC CELLS.
CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U76672; AAC49880.1; -
DR InterPro; IPR001019; Gprotein_alpha.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR SMART; SM00275; G-alpha; 1.
KW GTP-binding; Transducer; Multigene family.
FT NP_BIND 41 48 GTP (BY SIMILARITY).
FT NP_BIND 202 206 GTP (BY SIMILARITY).
FT NP_BIND 271 274 GTP (BY SIMILARITY).
SQ SEQUENCE 354 AA; 40405 MW; 6FB4D18305E76AC8 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 354;

Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0;

0; Indels 0; Gaps 0;

Qy 115 DPIIPAI 121

Db 135 DPIIPAI 141

Search completed: April 4, 2002, 14:50:10
Job time: 248 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 14:45:32 ; Search time 28.85 Seconds
(without alignments)
1754.254 Million cell updates/sec

Title: US-09-407-806A-4
Perfect score: 346
Sequence: 1 LRALVFHGNLQYAEIPKSEP.....ERRLDAFRAIYNDRNGNEP 346

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 473505 seqs, 146272329 residues
Word size : 0
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 18 | 5.2 | 364 | 1 Q9HBB5 | Q9hbb5 pyrococcus |
| 2 | 12 | 3.5 | 364 | 1 O58106 | O58106 pyrococcus |
| 3 | 8 | 2.3 | 102 | 2 Q9K2W4 | Q9k2w4 lactococcus |
| 4 | 8 | 2.3 | 105 | 2 Q9LC53 | Q9lc53 lactococcus |
| 5 | 8 | 2.3 | 218 | 1 Q58562 | Q58562 methanococc |
| 6 | 8 | 2.3 | 243 | 2 P73217 | P73217 synechocyst |
| 7 | 8 | 2.3 | 262 | 5 O16952 | O16952 caenorhabdi |
| 8 | 8 | 2.3 | 272 | 1 O59526 | O59526 pyrococcus |
| 9 | 8 | 2.3 | 272 | 1 Q9V1Z1 | Q9v1z1 pyrococcus |
| 10 | 8 | 2.3 | 328 | 5 O16953 | O16953 caenorhabdi |
| 11 | 8 | 2.3 | 466 | 2 O30418 | O30418 lactococcus |
| 12 | 8 | 2.3 | 466 | 2 Q9CG20 | Q9cg20 lactococcus |
| 13 | 8 | 2.3 | 492 | 10 O80497 | O80497 arabidopsis |
| 14 | 7 | 2.0 | 85 | 12 O90992 | O90992 human immun |
| 15 | 7 | 2.0 | 119 | 11 O08929 | O08929 mus musculu |
| 16 | 7 | 2.0 | 128 | 2 Q9CFK0 | Q9cfk0 lactococcus |
| 17 | 7 | 2.0 | 137 | 2 P96467 | P96467 streptococc |
| 18 | 7 | 2.0 | 147 | 9 O80190 | O80190 bacterioph |
| 19 | 7 | 2.0 | 160 | 10 Q9AVF7 | Q9avf7 amaranthus |

| | | | | | |
|----|---|-----|-----|-----------|--------------------|
| 20 | 7 | 2.0 | 175 | 1 O28131 | O28131 archaeoglob |
| 21 | 7 | 2.0 | 195 | 2 Q9R5Y6 | Q9r5y6 azotobacter |
| 22 | 7 | 2.0 | 196 | 11 Q9D702 | Q9d702 mus musculu |
| 23 | 7 | 2.0 | 212 | 12 Q9WH18 | Q9wh18 fish lympho |
| 24 | 7 | 2.0 | 215 | 11 O70315 | O70315 mus musculu |
| 25 | 7 | 2.0 | 221 | 2 Q9X4K5 | Q9x4k5 pseudomonas |
| 26 | 7 | 2.0 | 241 | 12 P89296 | P89296 yam mosaic |
| 27 | 7 | 2.0 | 242 | 6 Q9N1Q5 | Q9n1q5 hylobates l |
| 28 | 7 | 2.0 | 248 | 2 Q9KC51 | Q9kc51 bacillus ha |
| 29 | 7 | 2.0 | 252 | 10 Q9LGY4 | Q9lgy4 oryza sativ |
| 30 | 7 | 2.0 | 262 | 13 Q90713 | Q90713 gallus gall |
| 31 | 7 | 2.0 | 267 | 1 Q9V0E2 | Q9v0e2 pyrococcus |
| 32 | 7 | 2.0 | 271 | 2 Q9S4L4 | Q9s4l4 streptococc |
| 33 | 7 | 2.0 | 277 | 1 O58047 | O58047 pyrococcus |
| 34 | 7 | 2.0 | 277 | 2 Q9WXG7 | Q9wxg7 alcaligenes |
| 35 | 7 | 2.0 | 284 | 2 Q9ZHX1 | Q9zhx1 brucella me |
| 36 | 7 | 2.0 | 285 | 5 Q9U2H9 | Q9u2h9 caenorhabdi |
| 37 | 7 | 2.0 | 286 | 4 Q9NZ30 | Q9nz30 homo sapien |
| 38 | 7 | 2.0 | 291 | 4 Q9Y3A1 | Q9y3a1 homo sapien |
| 39 | 7 | 2.0 | 295 | 2 Q9CMB9 | Q9cmb9 pasteurella |
| 40 | 7 | 2.0 | 296 | 2 Q9A6J6 | Q9a6j6 caulobacter |
| 41 | 7 | 2.0 | 301 | 4 Q9NTX5 | Q9ntx5 homo sapien |
| 42 | 7 | 2.0 | 317 | 10 Q9AWS7 | Q9aws7 oryza sativ |
| 43 | 7 | 2.0 | 318 | 2 Q9PER5 | Q9per5 xylella fas |
| 44 | 7 | 2.0 | 325 | 4 Q9UFM6 | Q9ufm6 homo sapien |
| 45 | 7 | 2.0 | 325 | 4 Q9BTF9 | Q9btf9 homo sapien |

ALIGNMENTS

RESULT 1
Q9HBB5
ID Q9HBB5 PRELIMINARY; PRT; 364 AA.
AC Q9HBB5;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE ALPHA-GALACTOSIDASE.
GN GALA.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 3638;
RA Verhees C.H.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF195244; AAG28455.1;
SQ SEQUENCE 364 AA; 41545 MW; 3E8E0AF5BDCCF2A5 CRC64;

Query Match 5.2%; Score 18; DB 1; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RALVFHGNLQYAEIPKSE 19
|
Db 2 RALVFHGNLQYAEIPKSE 19

RESULT 2
O58106 PRELIMINARY; PRT; 364 AA.
ID O58106;
AC O58106;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE HYPOTHETICAL 41.8 KDA PROTEIN PH0368.
GN PH0368.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;

DB 70 EVIDELNS 77
 RESULT 6
 P73217 PRELIMINARY; PRT; 243 AA.
 AC P73217;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE HYPOTHETICAL 26.6 KDA PROTEIN.
 GN SLL1570.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D90904; BAA17244.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 243 AA; 26620 MW; 37DE109025773501 CRC64;
 Query Match 2.3%; Score 8; DB 2; Length 243;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 218 PKKVASWI 225
 DB 16 PKKVASWI 23
 RESULT 7
 O16952 PRELIMINARY; PRT; 262 AA.
 ID O16952;
 AC O16952;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE COSMID R11G11.
 GN R11G11.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilton R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailson N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RN Nature 368:32-38(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilton R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailson N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RN Nature 368:32-38(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilton R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailson N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RN Nature 368:32-38(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilton R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailson N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RN Nature 368:32-38(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilton R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailson N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RN Nature 368:32-38(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilton R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailson N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RN Nature 368:32-38(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilton R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailson N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RN Nature 368:32-38(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilton R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailson N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RN Nature 368:32-38(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilton R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Cop

GN PAB2165.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=ORSAY;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ248283; CAB49207.1; -.
DR InterPro; IPR002309; trna-synt_2.
DR Pfam; PF01336; trna_anti; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 272 AA; 31827 MW; 515DA586372DD837 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 272;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 254 LLEVIDEL 261
Db 173 LLEVIDEL 180

RESULT 10
O16953
ID O16953; PRELIMINARY; PRT; 328 AA.
AC O16953;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE COSMID R11G11.
GN R11G11.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Davidson S., Wohlmann P., Bauer C.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF022976; AAC69085.1; -.
SQ SEQUENCE 328 AA; 37192 MW; 12339872BA82C5D6 CRC64;

Query Match 2.3%; Score 8; DB 5; Length 328;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 LLLRELK 177
Db 140 LLLRELK 147
RESULT 11
O30418
ID O30418; PRELIMINARY; PRT; 466 AA.
AC O30418; O50645;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GLUTAMATE DECARBOXYLASE (EC 4.1.1.15).
GN GADB.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=MG1363;
RX MEDLINE=98143417; PubMed=9484886;
RA Sanders J.W., Leenhouts K., Burghoorn J., Brands J.R., Venema G.,
RA Kok J.;
RT "A chloride-inducible acid resistance mechanism in Lactococcus lactis
and its regulation.";
RL Mol. Microbiol. 27:299-310(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=01-7;
RA Nomura M., Fujita Y., Kobayashi M., Oomomo S.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CONVERTS INTERNALIZED GLUTAMATE TO GABA AND INCREASES
CC THE INTERNAL PH. INVOLVED IN GLUTAMATE-DEPENDENT ACID RESISTANCE.
CC -!- CATALYTIC ACTIVITY: L-GLUTAMATE = 4-AMINOBUTANOATE + CO(2).
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -!- INDUCTION: EXPRESSION IS HIGHEST AT ONSET OF STATIONARY PHASE IN
CC PRESENCE OF NaCl AND GLUTAMATE, AND AT LOW PH. CHLORIDE-DEPENDENT
CC EXPRESSION IS ACTIVATED BY GADR.
CC -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
DR EMBL; AF005098; AAC46188.1; -.
DR EMBL; AB010789; BAA24585.1; -.
DR InterPro; IPR002129; Pyridoxal_dec.
DR Pfam; PF00282; pyridoxal_dec; 1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
KW Lyase; Decarboxylase; Pyridoxal phosphate; Multigene family.
FT BINDING 277 277 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT VARIANT 114 114 G -> E (IN STRAIN 01-7).
FT VARIANT 116 117 IV -> FM (IN STRAIN 01-7).
FT VARIANT 172 172 V -> I (IN STRAIN 01-7).
FT VARIANT 185 185 R -> K (IN STRAIN 01-7).
FT VARIANT 193 193 E -> D (IN STRAIN 01-7).
FT VARIANT 353 353 Y -> F (IN STRAIN 01-7).
FT VARIANT 356 356 E -> K (IN STRAIN 01-7).
FT VARIANT 371 371 A -> S (IN STRAIN 01-7).
FT VARIANT 383 383 N -> D (IN STRAIN 01-7).
FT VARIANT 444 444 D -> E (IN STRAIN 01-7).
FT VARIANT 453 453 F -> Y (IN STRAIN 01-7).
FT VARIANT 455 455 Q -> E (IN STRAIN 01-7).
SQ SEQUENCE 466 AA; 53849 MW; 28B94EB3FEEB168D CRC64;

Query Match 2.3%; Score 8; DB 2; Length 466;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 DEGNARLN 302
Db 52 DEGNARLN 59

```
RESULT 12
Q9CG20
ID Q9CG20 PRELIMINARY; PRT; 466 AA.
AC Q9CG20;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE GLUTAMATE DECARBOXYLASE (EC 4.1.1.15).
GN GADB.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RA Bolotin A., Wincker P., Mauer S., Jaillon O., Malarne K.,
RA Weissbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis."
RL Genome Res. 0:0-0(2001).
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -!- SIMILARITY: PYRIDOXAL-DEPENDENT DECARBOXYLASE FAMILY.
DR EMBL; AE006361; AAK05388.1; -.
DR InterPro; IPR002129; Pyridoxal_deC.
DR Pfam; PF00282; Pyridoxal_deC; 1.
KW Complete proteome; Decarboxylase; Lyase; Pyridoxal phosphate.
SQ SEQUENCE 466 AA; 53926 MW; BCD732253E473C2 CRC64;

Query Match 2.3%; Score 8; DB 2; Length 466;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 DEGNARLN 302
Db 52 DEGNARLN 59

RESULT 13
Q80497
ID Q80497 PRELIMINARY; PRT; 492 AA.
AC Q80497;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE PUTATIVE GLYCOSYL HYDROLASE.
GN F16B22.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC F16B22 genomic sequence."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AC003672; AAC27459.1; -.
DR HSSP; P26221; 1TF4.
DR Mendel; 31158; Arath;1057;31158.
DR InterPro; IPR001701; Glyco_hydro_9.
DR Pfam; PF00759; Glyco_hydro_9; 1.
DR PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.
KW Hydrolase.
SQ SEQUENCE 492 AA; 53591 MW; 6FE816C89048D620 CRC64;
```

Query Match 2.3%; Score 8; DB 10; Length 492;

```
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 KLVFEGKV 187
Db 292 KLVFEGKV 299

RESULT 14
Q90992
ID Q90992 PRELIMINARY; PRT; 85 AA.
AC Q90992;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE GP120 (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VI1267;
RA Heyndrickx L., Janssens W., Coppens S., Vereecken K., Willems B.,
RA Franssen K., Colebunders R., Vandenbruaene M., Van Der Groen G.;
RT "HIV-1 C2V3 env diversity among Belgian individuals."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AJ228225; CAA12855.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR NON_TER 1
FT NON_TER 85
FT NON_TER 85
SQ SEQUENCE 85 AA; 9453 MW; 1D1C9745BBB8602C CRC64;

Query Match 2.0%; Score 7; DB 12; Length 85;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 DRVKEEL 100
Db 75 DRVKEEL 81

RESULT 15
Q08929
ID Q08929 PRELIMINARY; PRT; 119 AA.
AC Q08929;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CE14 PROTEIN (FRAGMENT).
GN CE14.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=LIVER;
RA Zimmermann W., Fischer B., Olsen A., Nedellec P., Thompson J.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBDJ databases.
DR EMBL; Y13561; CAA73896.1; -.
DR InterPro; IPR003599; Ig.
DR SMART; SM00409; IG; 1.
KW pregnancy.
FT NON_TER 1
FT NON_TER 119
FT NON_TER 119
SQ SEQUENCE 119 AA; 13635 MW; E4E6658159641235 CRC64;

Query Match 2.0%; Score 7; DB 11; Length 119;
Best Local Similarity 100.0%; Pred. No. 63;
```

| | | | | | | | | | |
|---------|-----|--------------|-----|------------|----|--------|----|------|----|
| Matches | 7; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
| QY | 270 | LKHSGRE | 276 | | | | | | |
| Db | 70 | LKHSGRE | 76 | | | | | | |

Search completed: April 4, 2002, 14:49:45
Job time: 253 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 14:34:16 ; Search time 25.56 Seconds
(without alignments)
1002.714 Million cell updates/sec

Title: US-09-407-806A-4
Perfect score: 1778
Sequence: 1 LRALVFHGNLQYAEIPKSEP.....ERRLDFAIRYNDWRNGNEP 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101: *
1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 1586 | 89.2 | 364 | AAW34643 | Thermostable alpha |
| 2 | 118.5 | 6.7 | 649 | AAR47504 | Pyrococcus furiosu |
| 3 | 100.5 | 5.7 | 653 | AAW54870 | Super heat resista |
| 4 | 97.5 | 5.5 | 434 | AAG82181 | S. epidermidis ope |
| 5 | 97.5 | 5.5 | 557 | AAG82196 | S. epidermidis ope |
| 6 | 97.5 | 5.5 | 1396 | AA136871 | Protein involved i |
| 7 | 93.5 | 5.3 | 216 | AAG29078 | Arabidopsis thalia |
| 8 | 93.5 | 5.3 | 218 | AAG29077 | Arabidopsis thalia |
| 9 | 93.5 | 5.3 | 235 | AA129076 | Arabidopsis thalia |
| 10 | 93 | 5.2 | 377 | AA134045 | Zea mays protein f |
| 11 | 93 | 5.2 | 465 | AAR88328 | Cotton UDP glucose |

| | | | | | | |
|----|------|-----|------|----|----------|--------------------|
| 12 | 93 | 5.2 | 473 | 21 | AA07943 | A uridine diphosph |
| 13 | 93 | 5.2 | 473 | 21 | AA07948 | A uridine diphosph |
| 14 | 93 | 5.2 | 473 | 21 | AAG34044 | Zea mays protein f |
| 15 | 91 | 5.1 | 335 | 21 | AA052610 | Helicobacter pylor |
| 16 | 91 | 5.1 | 480 | 21 | AA016318 | Pinus radiata UGP |
| 17 | 91 | 5.1 | 510 | 22 | AA046303 | H. pylori HPS068 p |
| 18 | 91 | 5.1 | 594 | 22 | AA043334 | Amino acid sequenc |
| 19 | 88.5 | 5.0 | 403 | 21 | AA05758 | Arabidopsis thalia |
| 20 | 88.5 | 5.0 | 403 | 21 | AA049835 | Arabidopsis thalia |
| 21 | 88.5 | 5.0 | 406 | 21 | AA05757 | Arabidopsis thalia |
| 22 | 88.5 | 5.0 | 406 | 21 | AA049834 | Arabidopsis thalia |
| 23 | 88.5 | 5.0 | 422 | 21 | AA05756 | Arabidopsis thalia |
| 24 | 88.5 | 5.0 | 422 | 21 | AA049833 | Arabidopsis thalia |
| 25 | 88.5 | 5.0 | 492 | 21 | AA021554 | Arabidopsis thalia |
| 26 | 87.5 | 4.9 | 919 | 22 | AA083181 | Corynebacterium th |
| 27 | 87.5 | 4.9 | 1118 | 22 | AA048264 | Rice magnesium che |
| 28 | 87 | 4.9 | 329 | 22 | AA082183 | S. epidermidis ope |
| 29 | 87 | 4.9 | 459 | 21 | AA041936 | Arabidopsis thalia |
| 30 | 87 | 4.9 | 908 | 20 | AA033450 | Chimeric Taq DNA p |
| 31 | 86.5 | 4.9 | 513 | 21 | AA074583 | Neisseria meningit |
| 32 | 86.5 | 4.9 | 527 | 22 | AA099962 | ERA binding domain |
| 33 | 86.5 | 4.9 | 1218 | 20 | AA034680 | Chlamydia pneumoni |
| 34 | 86 | 4.8 | 832 | 19 | AA023962 | Modified Taq polym |
| 35 | 85.5 | 4.8 | 286 | 21 | AA035951 | Human 3-hydroxyacy |
| 36 | 85.5 | 4.8 | 301 | 21 | AA087599 | Human fatty acid b |
| 37 | 85.5 | 4.8 | 474 | 21 | AA07947 | A uridine diphosph |
| 38 | 85.5 | 4.8 | 748 | 21 | AA051060 | Arabidopsis thalia |
| 39 | 85 | 4.8 | 534 | 21 | AA016317 | Eucalyptus grandis |
| 40 | 84.5 | 4.8 | 512 | 21 | AA074582 | Neisseria gonorrhe |
| 41 | 84.5 | 4.8 | 907 | 21 | AA036818 | Arabidopsis thalia |
| 42 | 84.5 | 4.8 | 912 | 21 | AA036817 | Arabidopsis thalia |
| 43 | 84 | 4.7 | 720 | 22 | AA062032 | Recombinant P. fur |
| 44 | 84 | 4.7 | 720 | 22 | AA062033 | P. furiosus helica |
| 45 | 84 | 4.7 | 832 | 19 | AA076205 | T. aquaticus DNA p |

ALIGNMENTS

| | |
|----------|--|
| RESULT | 1 |
| AAW34643 | |
| ID | AAW34643 standard; Protein; 364 AA. |
| XX | |
| AC | AAW34643; |
| XX | |
| DT | 27-MAR-1998 (first entry) |
| XX | |
| DE | Thermostable alpha-galactosidase AEDII12RA-alpha-gal-18GC. |
| XX | |
| KW | Alpha-galactosidase; alpha-glycosidase; thermostable enzyme; |
| KW | food processing; alpha glycoside hydrolysis; raffinose; |
| KW | stachyose; verbascose; bean; flatulence; AEDII12RA-alpha-gal-18GC. |
| XX | |
| OS | Thermococcus alcaliphilus strain AEDII12RA. |
| XX | |
| FH | Key Location/Qualifiers |
| FT | Misc-difference 329 |
| FT | /note= "encoded by CTT" |
| XX | |
| PN | WO9732974-A1. |
| XX | |
| PD | 12-SEP-1997. |
| XX | |
| PF | 05-FEB-1997; 97WO-US01452. |
| XX | |
| PR | 08-MAR-1996; 96US-0613220. |
| XX | |
| PA | (RECO-) RECOMBINANT BIOCATALYSIS INC. |
| XX | |
| PI | Murphy D, Reid J, Rudolph MJ; |
| XX | |
| DR | WPI; 1997-470541/43. |
| DR | N-PSDB; AAT93753. |

XX Pyrococcus sp.
OS JP10150986-A.
XX 09-JUN-1998.
XX 21-NOV-1996; 96JP-03111117.
XX 21-NOV-1996; 96JP-03111117.
XX (BEAB-) BE ABLE KK.
PA (NAGA-) NAGASE SEIKAGAKU KOGYO KK.
XX WPI; 1998-379989/33.
DR N-PSDB; AAV27026.
XX New 4-alpha-glucanotransferase which has been heat-treated - used at
PT high temperatures to transfer at least one glucose unit
XX Claim 4; Page 10-13; 18pp; Japanese.
XX The super heat-resistant 4-alpha-glucanotransferase has an optimum pH of
CC 6.0-8.0, with an optimum temperature at pH 7.5 of 100 deg. C. It has at
CC least 90% activity after heat-treatment at 100 deg. C for 30 minutes at
CC pH 7.5. It can transfer at least one glucose unit in alpha-1,4-glucan
CC to alpha-1,4-glucan by an alpha-1,4-glucoside bond.
XX Sequence 653 AA;

Query Match 5.7%; Score 100.5; DB 19; Length 653;
Best Local Similarity 19.3%; Pred. No. 0.33;
Matches 74; Conservative 67; Mismatches 127; Indels 115; Gaps 19;
QY 24 EKAYIPVIETLIKEP---FGLNITGYTLKEL----PKDIILVKGGIASDLIEIIGTS-Y 75
bB 28 ersyrpmetl-eeypnmkvavhysgpllebirdnkpehldlrslvkrqgleivvagfy 86
QY 76 TAILPLLLPSRVEAQVQDRVKE--ELFEVSPKGFWLPE-LADPIIPAILKDNNGYEYLF 132
Db 87 epvlasip--kedrivqieklkefarnlgyeargvwltervwqpelvkslraagidyviv 144
QY 133 DEAMLFSAHLNS-----AIKPIKPLPHLIKAQREKRFYISYLLRELR 176
Db 145 ddyhmsaglskdelfwpyytedggevityfpl-----dekrlryl--ipfrpvd 191
QY 177 KATKLVF-----EGKVTLKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWIEDK-- 228
Db 192 ktleyhlslddgdeskvavfhdgkfgvwp-----tyewvyekgw 233
QY 229 -----NILLYGTDIE-----FIGYRDIAG--RMSVEGL-----LEVID 259
Db 234 lreffdrvssderinlmlyseylqrfrprglvlpiasyfemsewslparqaklfvefve 293
QY 260 ELNSELCPSELKHSGRELYLRTSSWADKSLRIWREDEGNARLNMLYNMREGELAFLENSD 319
Db 294 elk-----kenkfdryrvfvggiwknffkypesnymhkrmlmvksa-----vrnnp 342
QY 320 ARGWPLPERRLDAFRAIYNDWRG 342
Db 343 arefilraqncday-----whg 359

RESULT 4
AAG82181
ID AAG82181 standard; Protein; 434 AA.
XX AAG82181;
XX 03-SEP-2001 (first entry)
DT S. epidermidis open reading frame protein sequence SEQ ID NO:1456.
XX DE

XX Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW vaccination; endocarditis.
XX Staphylococcus epidermidis.
OS WO200134809-A2.
XX 17-MAY-2001.
XX 09-NOV-2000; 2000WO-US30782.
XX 09-NOV-1999; 99US-0164258.
XX (GLAX) GLAXO GROUP LTD.
PA Kimmerly WJ;
XX WPI; 2001-316495/33.
DR N-PSDB; AAH53031.
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX Claim 18; Page 411; 2188pp; English.
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX Sequence 434 AA;

Query Match 5.5%; Score 97.5; DB 22; Length 434;
Best Local Similarity 22.4%; Pred. No. 0.37;
Matches 81; Conservative 53; Mismatches 142; Indels 85; Gaps 16;
QY 13 AEIPKSE-PKVIEWKAYIPVIETLIKEPFGNITGYTLKFLPKD---IILVKGGIASDLI 68
Db 6 agvkkserkkmlskqetlnkeplvkrd--glkgygyvveyrtdarltevmkkaeaga 63
QY 69 EIIGTSYTAILPLPLSRVEAQVQDRVKEELFEVSPK-----GFWLPELADPIIPAIL 122
Db 64 eiinytksehftysdnkkvngievldmidgetyaikakkvinasgppwvdev----- 114
QY 123 KONGYEYLFADAEMLFSAHLNSAIKPIK-PLPHLIKAQREKRFYISYLLRE----- 174
Db 115 --rsgdyarnnkqlrltkgvhvvidqskfplgqavfydtekgrmi-faipregkayvgt 171
QY 175 -----LRKATKLVFEKGKVTLKVKDIEAVPVVAVNTAVMLIGRL 213
Db 172 tdtfydnekatplttqedrdylinainymfp---tvnvkdediestwagirplilekgd 228
QY 214 PLMNPKKVASWIEDKNILLYGTDIEFIGYRDIAGRMSVEGLEVDENSELCPSELKHS 273
Db 229 pseisrkddevw-egesglitagiagkltgyrhma-----leivdll-----akrl 271

nongonococcal urethritis, epididymitis, cervicitis, salpingitis, perihepatitis, bartholinitis; pneumopathy in breast feeding infants; and venereal lymphogranulomatosis. The polypeptides of the invention may be of use in treating these diseases.

| Sequence | 1396 AA; |
|----------|----------|
| SQ | |

| | | | | |
|-----------------------|------------------|-----------------|------------|--------------|
| Query Match | 5.5%; | Score 97.5; | DB 20; | Length 1396; |
| Best Local Similarity | 22.0%; | Pred. No. 2; | | |
| Matches 83; | Conservative 61; | Mismatches 144; | Indels 89; | Gaps 18; |

| | | | |
|----|-----|--|-----|
| QY | 15 | IPKSEPKVIEKAYIPVIETLKEEPPGLN--ITGVTLKFLPKDIIILVKGGIASDLIEIG | 72 |
| Dd | 404 | iqrqapevwd-----vleeliikhgvllnraptlhrigiqafepvliegk-airvhplvc | 457 |
| QY | 73 | TSYTA-----ILPLLRSRVEAQVRDRVKEELFEVSPKGFWLPPELADPI-IPAILKDN | 125 |
| Dd | 458 | aafnadfgdmavhvpls-ieaqle----akvlmmapdniflpssgkpvatpskdmrtl | 511 |
| QY | 126 | GYEYLFADAEMLFSAHLNSAIKPICKPLHLIKA-----QREKRFRYISYL | 170 |
| Dd | 512 | giyylmadtyfpeeh-ggktkfkdevevrlralnaggfilkdeicgsrrdegrgi--- | 567 |
| QY | 171 | LLRELRLKAIVFEGKVTLKVKDIEAVPVWVAVENTAV-----MLIGRLPLMN | 217 |
| Dd | 568 | ---hihekikvrldggi-----iettpgrvfntivpkelgfqnysmpskriselilqc | 618 |
| QY | 218 | PKKVASWIEDKNILLYGTDIEFIGYVDIAGRMVSVEGLEEV-IDELNSELCPSLUKHSGRE | 276 |
| Dd | 619 | ykkvgl---eatrvrflddlkelgfvqstkaaismglkvdkipeikkei---lkdaydk | 670 |
| QY | 277 | LYLRTSSWAD-----KSRLTWREDEGNARLNLMly-----NMRGELAFIAENSD | 319 |
| Dd | 671 | vaiuvkkqyedgitdgerhshtisiwte-vsdllsnalyseikkqtlnskhnoplflmvxsq | 729 |
| QY | 320 | ARGWPLPERRLDAFRAI | 336 |
| Dd | 730 | arqnksqlkqialrqrl | 746 |

RESULT 7
AAG29078
ID AAG29078 standard; Protein: 216 AA.
XX
AC AAG29078;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 34533.
XX
KW protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

| | |
|----|------------------------------|
| OS | Arabidopsis thaliana. |
| XX | |
| PN | EP1033405-A2. |
| XX | |
| PD | 06-SEP-2000. |
| XX | |
| PF | 25-FEB-2000; 2000EP-0301439. |
| XX | |
| PR | 25-FEB-1999; 99US-0121825. |
| PR | 05-MAR-1999; 99US-0123180. |
| PR | 09-MAR-1999; 99US-0123548. |
| PR | 23-MAR-1999; 99US-0125788. |
| PR | 25-MAR-1999; 99US-0126264. |
| PR | 29-MAR-1999; 99US-0126785. |
| PR | 01-APR-1999; 99US-0127462. |
| PR | 06-APR-1999; 99US-0128234. |
| PR | 08-APR-1999; 99US-0128714. |
| PR | 16-APR-1999; 99US-0129845. |

| | | |
|----|--------------|---------------|
| PR | 29-MAR-1999; | 99US-0126785. |
| PR | 01-APR-1999; | 99US-0127462. |
| PR | 06-APR-1999; | 99US-0128234. |
| PR | 08-APR-1999; | 99US-0128714. |
| PR | 16-APR-1999; | 99US-0129845. |
| PR | 19-APR-1999; | 99US-0130077. |
| PR | 21-APR-1999; | 99US-0130449. |
| PR | 23-APR-1999; | 99US-0130510. |
| PR | 23-APR-1999; | 99US-0130891. |
| PR | 28-APR-1999; | 99US-0131449. |
| PR | 30-APR-1999; | 99US-0132048. |
| PR | 30-APR-1999; | 99US-0132407. |
| PR | 04-MAY-1999; | 99US-0132484. |
| PR | 05-MAY-1999; | 99US-0132485. |
| PR | 06-MAY-1999; | 99US-0132486. |
| PR | 06-MAY-1999; | 99US-0132487. |
| PR | 07-MAY-1999; | 99US-0132863. |
| PR | 11-MAY-1999; | 99US-0134256. |
| PR | 14-MAY-1999; | 99US-0134218. |
| PR | 14-MAY-1999; | 99US-0134219. |
| PR | 14-MAY-1999; | 99US-0134221. |
| PR | 14-MAY-1999; | 99US-0134370. |
| PR | 18-MAY-1999; | 99US-0134768. |
| PR | 19-MAY-1999; | 99US-0134941. |
| PR | 20-MAY-1999; | 99US-0135124. |
| PR | 21-MAY-1999; | 99US-0135353. |
| PR | 24-MAY-1999; | 99US-0135629. |
| PR | 25-MAY-1999; | 99US-0136021. |
| PR | 27-MAY-1999; | 99US-0136392. |
| PR | 28-MAY-1999; | 99US-0136782. |
| PR | 01-JUN-1999; | 99US-0137222. |
| PR | 03-JUN-1999; | 99US-0137528. |
| PR | 04-JUN-1999; | 99US-0137502. |
| PR | 07-JUN-1999; | 99US-0137724. |
| PR | 08-JUN-1999; | 99US-0138094. |
| PR | 10-JUN-1999; | 99US-0138540. |
| PR | 10-JUN-1999; | 99US-0138847. |
| PR | 14-JUN-1999; | 99US-0139119. |
| PR | 16-JUN-1999; | 99US-0139452. |
| PR | 16-JUN-1999; | 99US-0139453. |
| PR | 17-JUN-1999; | 99US-0139492. |
| PR | 18-JUN-1999; | 99US-0139454. |
| PR | 18-JUN-1999; | 99US-0139455. |
| PR | 18-JUN-1999; | 99US-0139456. |
| PR | 18-JUN-1999; | 99US-0139457. |
| PR | 18-JUN-1999; | 99US-0139458. |
| PR | 18-JUN-1999; | 99US-0139459. |
| PR | 18-JUN-1999; | 99US-0139460. |
| PR | 18-JUN-1999; | 99US-0139461. |
| PR | 18-JUN-1999; | 99US-0139462. |
| PR | 18-JUN-1999; | 99US-0139463. |
| PR | 18-JUN-1999; | 99US-0139750. |
| PR | 18-JUN-1999; | 99US-0139763. |
| PR | 21-JUN-1999; | 99US-0139817. |
| PR | 22-JUN-1999; | 99US-0139899. |
| PR | 23-JUN-1999; | 99US-0140353. |
| PR | 23-JUN-1999; | 99US-0140354. |
| PR | 24-JUN-1999; | 99US-0140695. |
| PR | 28-JUN-1999; | 99US-0140823. |
| PR | 29-JUN-1999; | 99US-0140991. |
| PR | 30-JUN-1999; | 99US-0141287. |
| PR | 01-JUL-1999; | 99US-0141842. |
| PR | 01-JUL-1999; | 99US-0142154. |
| PR | 02-JUL-1999; | 99US-0142055. |
| PR | 06-JUL-1999; | 99US-0142390. |
| PR | 08-JUL-1999; | 99US-0142803. |
| PR | 09-JUL-1999; | 99US-0142920. |
| PR | 12-JUL-1999; | 99US-0142977. |
| PR | 13-JUL-1999; | 99US-0143542. |
| PR | 14-JUL-1999; | 99US-0143624. |
| PR | 15-JUL-1999; | 99US-0144005. |
| PR | 16-JUL-1999; | 99US-0144085. |
| PR | 16-JUL-1999; | 99US-0144086. |
| PR | 19-JUL-1999; | 99US-0144325. |
| PR | 19-JUL-1999; | 99US-0144331. |
| PR | 19-JUL-1999; | 99US-0144332. |
| PR | 19-JUL-1999; | 99US-0144333. |
| PR | 19-JUL-1999; | 99US-0144334. |
| PR | 19-JUL-1999; | 99US-0144335. |
| PR | 20-JUL-1999; | 99US-0144352. |
| PR | 20-JUL-1999; | 99US-0144632. |
| PR | 20-JUL-1999; | 99US-0144884. |
| PR | 21-JUL-1999; | 99US-0144814. |
| PR | 21-JUL-1999; | 99US-0145086. |
| PR | 21-JUL-1999; | 99US-0145088. |
| PR | 22-JUL-1999; | 99US-0145085. |
| PR | 22-JUL-1999; | 99US-0145087. |
| PR | 22-JUL-1999; | 99US-0145089. |
| PR | 22-JUL-1999; | 99US-0145192. |
| PR | 23-JUL-1999; | 99US-0145145. |
| PR | 23-JUL-1999; | 99US-0145218. |
| PR | 23-JUL-1999; | 99US-0145224. |
| PR | 26-JUL-1999; | 99US-0145276. |
| PR | 27-JUL-1999; | 99US-0145913. |
| PR | 27-JUL-1999; | 99US-0145918. |
| PR | 27-JUL-1999; | 99US-0145919. |
| PR | 28-JUL-1999; | 99US-0145951. |
| PR | 02-AUG-1999; | 99US-0146386. |
| PR | 02-AUG-1999; | 99US-0146388. |
| PR | 02-AUG-1999; | 99US-0146389. |
| PR | 03-AUG-1999; | 99US-0147038. |
| PR | 04-AUG-1999; | 99US-0147204. |
| PR | 04-AUG-1999; | 99US-0147302. |
| PR | 05-AUG-1999; | 99US-0147192. |
| PR | 05-AUG-1999; | 99US-0147260. |
| PR | 06-AUG-1999; | 99US-0147303. |
| PR | 06-AUG-1999; | 99US-0147416. |
| PR | 09-AUG-1999; | 99US-0147493. |
| PR | 09-AUG-1999; | 99US-0147935. |
| PR | 10-AUG-1999; | 99US-0148171. |
| PR | 11-AUG-1999; | 99US-0148319. |
| PR | 12-AUG-1999; | 99US-0148341. |
| PR | 13-AUG-1999; | 99US-0148565. |
| PR | 13-AUG-1999; | 99US-0148684. |
| PR | 16-AUG-1999; | 99US-0149368. |
| PR | 17-AUG-1999; | 99US-0149175. |
| PR | 18-AUG-1999; | 99US-0149426. |
| PR | 20-AUG-1999; | 99US-0149722. |
| PR | 20-AUG-1999; | 99US-0149723. |
| PR | 20-AUG-1999; | 99US-0149929. |
| PR | 23-AUG-1999; | 99US-0149902. |
| PR | 23-AUG-1999; | 99US-0149930. |
| PR | 25-AUG-1999; | 99US-0150566. |
| PR | 26-AUG-1999; | 99US-0150884. |
| PR | 27-AUG-1999; | 99US-0151065. |
| PR | 27-AUG-1999; | 99US-0151066. |
| PR | 27-AUG-1999; | 99US-0151080. |
| PR | 30-AUG-1999; | 99US-0151303. |
| PR | 31-AUG-1999; | 99US-0151438. |
| PR | 01-SEP-1999; | 99US-0151930. |
| PR | 07-SEP-1999; | 99US-0152363. |
| PR | 10-SEP-1999; | 99US-0153070. |
| PR | 13-SEP-1999; | 99US-0153758. |
| PR | 15-SEP-1999; | 99US-0154018. |
| PR | 16-SEP-1999; | 99US-0154039. |
| PR | 20-SEP-1999; | 99US-0154779. |
| PR | 22-SEP-1999; | 99US-0155139. |
| PR | 23-SEP-1999; | 99US-0155486. |
| PR | 24-SEP-1999; | 99US-0155659. |
| PR | 28-SEP-1999; | 99US-0156458. |
| PR | 29-SEP-1999; | 99US-0156596. |
| PR | 04-OCT-1999; | 99US-0157117. |
| PR | 05-OCT-1999; | 99US-0157753. |
| PR | 06-OCT-1999; | 99US-0157865. |
| PR | 07-OCT-1999; | 99US-0158029. |
| PR | 08-OCT-1999; | 99US-0158232. |

09-FEB-2000; 2000WO-US03513.

PF 09-FEB-2000; 2000WO-US03513.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.2%; Score 93; DB 21; Length 473;
Best Local Similarity 25.3%; Pred. No. 1.2;

Matches 59; Conservative 24; Mismatches 80; Indels 70; Gaps 12;

QY 53 PKDIILVKGGA-SDL---IEIIGTSYTAILPLPL-----SRVEA 89
Db 102 pksvievrngftfldlivigieslnkkygcnvplllmnsfnthddtqkivekysnsniei 161
QY 90 ----QVQRDRVKEELFEVSP-----KGFWLPADPIIPA-----ILKNGYEYLF 131
Db 162 htfnqgyprivtedfplpksksgkdgyppghgdvfpnlsgkldillaqgkeyvf 221
QY 132 ADEAMLFSAHLNSAIKPIKPLPHLIKAQREKFRFYISVLLRELRLKAIKLVFEKVTL-- 189
Db 222 van-----sdnl-gaivdikilnhlinnqnecyme-vtpktladvkgtllisyegrvqlle 275
QY 190 -----KVKDIEAVPV-----WVAVNTAVMLIGRLPLM-----NPKKV 221
Db 276 iaqvpdehvnfeksiekfifntnnlwnlkavkriveaealkmeiipnpkev 328

RESULT 15
AAB52610
ID AAB52610 standard; Protein; 335 AA.

AC AAB52610;

DT 23-FEB-2001 (first entry)

DE Helicobacter pylori bait polypeptide #128.

XX Helicobacter pylori; two-hybrid system; protein-protein interaction;
KW bait polypeptide; gastric ulcer; antibacterial.

OS Helicobacter pylori.

XX WO200066722-A1.

PN 09-NOV-2000.

PD 14-APR-2000; 2000WO-IB00603.

XX PF

XX 30-APR-1999; 99EP-0401066.
PR (HYBR-) HYBRIGENICS SA.
XX Legrain P, Selig L, Rain J;
XX WPI; 2000-687535/67.
DR N-PSDB; AAC97356.

XX A two-hybrid system for identifying compounds useful in the treatment
PT of e.g. gastric ulcers comprises producing a collection of recombinant
PT cell clones -

XX Example 5; Page 242-243; 267pp; English.

XX The present sequence is a bait polypeptide used in a Helicobacter
CC pylori two-hybrid screen to identify protein-protein interactions.
CC The method is used to identify a recombinant cell clone expressing a
CC prey polypeptide which is capable of interacting with the bait
CC polypeptide. The two hybrid system is useful for screening compounds
CC for antibacterial activity. It may be used in the treatment of gastric
CC ulcers. The polynucleotides are useful as amplification primers or
CC specific detection probes. The polypeptides, vectors or host cells can
CC be used as immunogens to produce mono- or polyclonal antibodies. The
CC polynucleotides, polypeptides, antibodies, vectors, host cells or
CC modulating agents can be used to produce a pharmaceutical composition.

XX Sequence 335 AA;

Query Match 5.1%; Score 91; DB 21; Length 335;
Best Local Similarity 21.2%; Pred. No. 1.1;
Matches 47; Conservative 46; Mismatches 71; Indels 58; Gaps 10;

QY 5 VEHGNLQYAEIPKSEPKVIEKAYIPVETLKEEPFG----LNITGYTLKFLPKDIILVK 60
Db 44 vyygmvlnaevaetleev-ekghykhfqnalqmkgvgqiarvetlgaqvaydkahiasvk 102
QY 61 GGIASDLIEIIGTSYTAIL----PLPLSRVEAQVQRDRVKEELFEVSPKGFWLPADP 116
Db 103 ---akdvlevsqslsfnsilsskddlvpskkleirtekn-----lpdisff 144
QY 117 IIPAILKNGYEYLFADAMLFSAHLNSAIKPIKPLPHLIKAQREKFRFYISVLLRELRL 176
Db 145 vssti---nsypvlktlenqiqiskentklqiakflpqv-----sffgsyim---- 188
QY 177 KAIKLVFEKGVTLKVKDIEAVPVVAVNTAVMLIGRLPLMNP 218
Db 189 kqnnsvfe-----dmipswf-----vgvagrmpilsp 215

Search completed: April 4, 2002, 14:40:13
Job time: 357 sec

QY 1 LRALVFHGNLQYAEIPKSEPKVIEKAYIPVIETLIKEEPPFGLNITGYTLKFLPKDIILVK 60
Db 1 LRALVFHGNLQYAEIPKSEPKVIEKAYIPVIETLIKEEPPFGLNITGYTLKFLPKDIILVK 60
QY 61 GGIASDLIEIIGTSYTAILPLPLSRVEAQVQRDRVKEELFEVSPKGFWLPELADPIIPA 120
Db 61 GGIASDLIEIIGTSYTAILPLPLSRVEAQVQRDRVKEELFEVSPKGFWLPELADPIIPA 120
QY 121 ILKNGYEYLFADAEAMLFSAHLNSAIKPIKPLPHLIKAQREKFRFYISYLLRELKAIK 180
Db 121 ILKNGYEYLFADAEAMLFSAHLNSAIKPIKPLPHLIKAQREKFRFYISYLLRELKAIK 180
QY 181 LVFEGKVTLKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWIEDKNILLYGTDIEFI 240
Db 181 LVFEGKVTLKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWIEDKNILLYGTDIEFI 240
QY 241 GYRDIAGRMSVEGLLEVIDELNSELCPSELKHSGRELYLRTSSWADKSLRIWREDEGNAR 300
Db 241 GYRDIAGRMSVEGLLEVIDELNSELCPSELKHSGRELYLRTSSWADKSLRIWREDEGNAR 300
QY 301 LNMLYNMRGELAFLAENSDARGWPLPERRLDARAIYNDWRNGNEP 346
Db 301 LNMLYNMRGELAFLAENSDARGWPLPERRLDARAIYNDWRNGNEP 346

RESULT 2
US-07-894-212A-8
; Sequence 8, Application US/07894212A
; Patent No. 5366883
; GENERAL INFORMATION:
; APPLICANT: ASADA, KIYOZO
; APPLICANT: UEMORI, TAKASHI
; APPLICANT: MUKAI, HIROYUKI
; APPLICANT: KATO, IKUNOSHIN
; APPLICANT: LADERMAN, KENNETH
; APPLICANT: ANFINSEN, CHRISTIAN
; TITLE OF INVENTION: THE ALPHA-AMYLASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/894,212A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 95469/C-1195
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 647 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-894-212A-8

Query Match 6.6%; Score 116.5; DB 1; Length 647;
Best Local Similarity 20.2%; Pred. No. 0.0018;

Matches 79; Conservative 64; Mismatches 117; Indels 131; Gaps 21;
QY 24 EKAYIPVIETLIKEEPPFGLNITGYTLKFL----PKDIILVKGGIASDLIEIIGTS-Y 75
Db 27 EKCYWPFLETL-EEYPNMKVAIHTSGPLIEWLQDNRPYIDLLRSLVKRGQVEIVVAGFY 85
QY 76 TAILPLPLSRVEAQVQRDRVKEELFE--VSPKGFWLPE-LADPIIPAILKDNNGYEYLF 131
Db 86 EPVLASIP--KEDRIEQIRLMKEWAKSIGFDARGVWLTERVWQPELVKTLKESGIDYVI 142
QY 132 ADEAMLFSAHLNS-----AIKPIKPLPHLIKAQREKFRFYISYLLREL 175
Db 143 VDDYHFMASAGLSKEELYWPYYTDDGGGEVIAVFPI-----DEKLRYL--IPFRPV 189
QY 176 RKAIKLVF-----EGKVTLKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWIEDKN 229
Db 190 DKVLEYLHSLIDGDESKVAVFHDDGKFGIWPG-----TYEWVYEKG 231
QY 230 ILLYGTDIEFIGYRDIAGRMSVEGLLEVIDELNSELCPSEL-KHSGREL-----YLRT 281
Db 232 WL-----REFFDRISSD-----EKINLMLYTEYLEKYKPRGLVYLPIASYFEM 274
QY 282 SSWA--DKSLRIWREDEGNARLNMLYN-----MRGEL-----AF 313
Db 275 SEWSLPAKQARLFVEFVNLKVGIFEKYRVFVRCGIWKGNFFYKYPESNYMHKRMMLVSK 334
QY 314 LAENSDARGWPLPERRLDARAIYND--WRG 342
Db 335 LVRNN-----PEARKYLLRAQCNDAYWHG 358

RESULT 3
US-07-894-212A-2
; Sequence 2, Application US/07894212A
; Patent No. 5366883
; GENERAL INFORMATION:
; APPLICANT: ASADA, KIYOZO
; APPLICANT: UEMORI, TAKASHI
; APPLICANT: MUKAI, HIROYUKI
; APPLICANT: KATO, IKUNOSHIN
; APPLICANT: LADERMAN, KENNETH
; APPLICANT: ANFINSEN, CHRISTIAN
; TITLE OF INVENTION: THE ALPHA-AMYLASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/894,212A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 95469/C-1195
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.


```

QY      81 LLPLSRVEAQVQRDRVK-EELFEVSQGFGLPELADPIIPA-----ILKDN 12
Db      241 CIIIVHDPNLN-NSDKMKFNTYYLLLEYKEYW-HQLWSQIIPAHQT VKIQERTGISEVVQNS 298
QY      126 GYE----YLFADEAMLFS AHLNSAIKPI-----KPLPHLIKAOREK-----RF 164
Db      299 MIEDLNYIGADFGMLFYFRSSGFKEQITRGLNRPLSQTTTQLGERVEEMEYNSNDLDV 358
QY      165 RYISYLLLRELKRAIKLVFEGKVTLKKVDIEAVPVWVAVN---TAVMLIGRLPLMNPKKV 221
Db      359 RVKYALAREF-----TLKRVRNGEIVKNWVAVDYRLAGIQSYNPAPITNP--- 403
QY      222 ASWIEDKNILLYGTDIEFIGYRDIAGRMSVEGLL-----EVIDELNSELCP 267
Db      404 -----LTLTKHTTIIRCENSYDGHIFKTPLIFKNGEIVIVKTNEELIP 444

RESULT 11
US-08-788-892-2
; Sequence 2, Application US/08788892
; Patent No. 5989864
; GENERAL INFORMATION:
; APPLICANT: Burnham, Martin
; APPLICANT: Gentry, Daniel
; APPLICANT: Hodgson, John
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 5989864el spo-rel
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,892
; FILING DATE: 22-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9604045.6
; FILING DATE: 26-FEB-1996
; APPLICATION NUMBER: 60/029329
; FILING DATE: 29-OCT-1996
; APPLICATION NUMBER: 60/030151
; FILING DATE: 01-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: GM50002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 736 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-788-892-2

```


Query Match 4.6%; Score 82; DB 1; Length 893;
Best Local Similarity 22.4%; Pred. No. 11;
Matches 85; Conservative 49; Mismatches 103; Indels 142; Gaps 22;

QY 19 EPKVIKAYIP-----VIETL--IKEEPFGLNITGYTLKFLPKDIILVKGGI 63
:||| | ||| :||| | ||| | ||| | ||| | ||| | ||| :||| |
Db 344 KPK--EAYIPLHHRNAQNLDKEVLKKEILEDP-GAKIVGNLKFYDK-VLMVKG-- 397
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
QY 64 ASDLIEIIGTSYTAILPLPLSRVEAQVDRVKEELFEVSPKGFWLPELADPIIPAILK 123
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
Db 398 -----VEPVPPYFDT-----MIAAYLLEPNEKKFNLDLLA-----LK 429
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
QY 124 DNGYEYLFADAEMLFSAHL---NSAIKPIKPLPHLIKAQREKFRY-----ISYLLRE 174
:||| | ||| | ||| | ||| :||| | ||| :||| | ||| :||| |
Db 430 FLGYKMTSYQELMSFSFPLFGFSFADVPV-----EKAANYSCEDADITYRLYKT 478
:||| | ||| | ||| | ||| :||| | ||| :||| | ||| :||| |
QY 175 LRKAIKLVFEGKVTLKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWIEDKNILL-- 232
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
Db 479 L-----SLKLHEADLENVFIKI-----EMPLVN---VLARMELNGVYVDT 515
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
QY 233 -----YGTDIEFIG---YRDIAG-----RMSVEGL----- 254
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
Db 516 EFLKKLSEYGGKLEELAEIYR-IAGEPFNINSPKQVSRILFEKLGKPRGKTKTG DY 574
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
QY 255 ---LEVIDEL--NSELCPSELKHSGRELYLRTSSWADKSLRIWREDEGNARLNMLYNMRG 309
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
Db 575 STRIEVLEELAGEHEIIPILLEY--RKIQKSTYIDALPKMVNPKTG--RIHASFNQTG 630
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
QY 310 ELAFLAENS DARGWPLPER 328
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
Db 631 TATGRLLSSDDPNLQNLPTK 649
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |

RESULT 14
US-09-105-697-10
; Sequence 10, Application US/09105697
; Patent No. 6228628
; GENERAL INFORMATION:
; APPLICANT: Gelfand Ph.D., David H.
; APPLICANT: Reichert, Fred L.
; TITLE OF INVENTION: MUTANT CHIMERIC DNA POLYMERASE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Roche Molecular Systems
; STREET: 1080 U.S. Highway 202
; CITY: Branchburg
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 08876
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,697
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Petry Ph.D., Douglas A.
; REGISTRATION NUMBER: 35321
; REFERENCE/DOCKET NUMBER: 1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510)814-2974
; TELEFAX: (510)814-2977
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 893 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-105-697-10

Query Match 4.6%; Score 82; DB 4; Length 893;
Best Local Similarity 22.4%; Pred. No. 11;
Matches 85; Conservative 49; Mismatches 103; Indels 142; Gaps 22;

QY 19 EPKVIKAYIP-----VIETL--IKEEPFGLNITGYTLKFLPKDIILVKGGI 63
:||| | ||| :||| | ||| | ||| | ||| | ||| | ||| :||| |
Db 344 KPK--EAYIPLHHRNAQNLDKEVLKKEILEDP-GAKIVGNLKFYDK-VLMVKG-- 397
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
QY 64 ASDLIEIIGTSYTAILPLPLSRVEAQVDRVKEELFEVSPKGFWLPELADPIIPAILK 123
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
Db 398 -----VEPVPPYFDT-----MIAAYLLEPNEKKFNLDLLA-----LK 429
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
QY 124 DNGYEYLFADAEMLFSAHL---NSAIKPIKPLPHLIKAQREKFRY-----ISYLLRE 174
:||| | ||| | ||| | ||| :||| | ||| :||| | ||| :||| |
Db 430 FLGYKMTSYQELMSFSFPLFGFSFADVPV-----EKAANYSCEDADITYRLYKT 478
:||| | ||| | ||| | ||| :||| | ||| :||| | ||| :||| |
QY 175 LRKAIKLVFEGKVTLKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWIEDKNILL-- 232
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
Db 479 L-----SLKLHEADLENVFIKI-----EMPLVN---VLARMELNGVYVDT 515
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
QY 233 -----YGTDIEFIG---YRDIAG-----RMSVEGL----- 254
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
Db 516 EFLKKLSEYGGKLEELAEIYR-IAGEPFNINSPKQVSRILFEKLGKPRGKTKTG DY 574
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
QY 255 ---LEVIDEL--NSELCPSELKHSGRELYLRTSSWADKSLRIWREDEGNARLNMLYNMRG 309
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
Db 575 STRIEVLEELAGEHEIIPILLEY--RKIQKSTYIDALPKMVNPKTG--RIHASFNQTG 630
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
QY 310 ELAFLAENS DARGWPLPER 328
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |
Db 631 TATGRLLSSDDPNLQNLPTK 649
:||| | ||| :||| | ||| :||| | ||| :||| | ||| :||| |

RESULT 15
PCT-US91-07035-4
; Sequence 4, Application PC/TUS9107035
; GENERAL INFORMATION:
; APPLICANT: Gelfand, David H.
; APPLICANT: Abramson, Richard D.
; TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
; TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cetus Corporation
; STREET: 1400 Fifty-third Street
; CITY: Emeryville
; STATE: California
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/07035
; FILING DATE: 19910930
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 590,490
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 590,466
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 590,213
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 523,394
; FILING DATE: 15-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 143,441

Search completed: April 4, 2002, 14:40:44
Job time: 193 sec

| | | | | | | |
|----|-----------------------|---|---------------|-------|-------------|---------------------------|
| | Query Match | 4.6%; | Score 82; | DB 5; | Length 893; | |
| | Best Local Similarity | 22.4%; | Pred. No. 11; | | | |
| | Matches | 85; | Conservative | 49; | Mismatches | 103; Indels 142; Gaps 22; |
| QY | 19 | EPKVIEKAYIP-----VIETL--IKEEPFGLNITGYTLKFLPKDIIIVKGGI | 63 | | | |
| Db | 344 | KPK--EAYIPLHHRNAQNLDKEVKKLKEILEDP-GAKIVGQNLFKDYK-VLMVKG-- | 397 | | | |
| QY | 64 | ASDLIEIGTSYTAIPLLLPLSRVEAQVRDRVKEELFEVSPKGFWLPPELADPIIPATLK | 123 | | | |
| Db | 398 | -----VEPVPPYFDT-----MIAAYILLEPNEKKFNLDLA-----LK | 429 | | | |
| QY | 124 | DNGYEYLFADAEMLFSAHL---NSAIKPIKLPHLIKAQREKFRY-----ISYLLLE | 174 | | | |
| Db | 430 | FLGYKWTYSQELMSFSFPFLFGSFADV-----EKAANYSCEDADITYRLYKT | 478 | | | |
| QY | 175 | LRKAIKLVFEGKYTLKVVDIEAVPVWVAVENTAVMLIGRLPLMNPKKVASWIEDKNILL-- | 232 | | | |
| Db | 479 | L-----SLKLHEADLENVFYKI-----EMPLVN--VLARMELNGVVYDT | 515 | | | |
| QY | 233 | -----YGTDIEFIG--YRDIAG-----RMSVEGL----- | 254 | | | |
| Db | 516 | EFLKKLSEYGGKKLEEAAEEIYR-IAGEPFNINSPKQVSRIILFEKLGIKPRGKTTKTGDY | 574 | | | |
| QY | 255 | ---LEVIDEL--NSELCPSSELKHSGRELYLRTSSWADKSLRIWDEGNARLNMLYNMRG | 309 | | | |
| Db | 575 | STRIEVLEELAGEHEIIPILLEY--RKIQKLKSTYIDALPKMVNPKTG--RIHASFNQTG | 630 | | | |
| QY | 310 | ELAFLAENSDARGWP LPER | 328 | | | |
| Db | 631 | TATGRLLSSDDPNLQNLP TK | 649 | | | |

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 14:38:16 ; Search time 21.67 Seconds
(without alignments)
1216.262 Million cell updates/sec

Title: US-09-407-806A-4
Perfect score: 1778
Sequence: 1 LRALVFHGNLQYAEIPKSEP.....ERRLDAFRAIYNDRNGNEP 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 1229 | 69.1 | 364 | 2 E71144 | hypothetical prote |
| 2 | 151 | 8.5 | 686 | 1 ALDYAT | amylase A (EC 3.2. |
| 3 | 126.5 | 7.1 | 633 | 2 G71241 | probable alpha-am |
| 4 | 120.5 | 6.8 | 655 | 2 E75206 | alpha-amylase (or |
| 5 | 116.5 | 6.6 | 649 | 2 A49512 | alpha-amylase (EC |
| 6 | 111 | 6.2 | 4725 | 1 A44357 | dynein heavy chain |
| 7 | 110 | 6.2 | 1362 | 2 A75207 | amylopullulanase P |
| 8 | 106.5 | 6.0 | 392 | 2 H82815 | fimbrial assembly |
| 9 | 106.5 | 6.0 | 467 | 2 B64501 | alpha-amylase (EC |
| 10 | 104.5 | 5.9 | 1251 | 2 JH0256 | botulinum neurotox |
| 11 | 102 | 5.7 | 738 | 2 S52737 | NADH dehydrogenase |
| 12 | 102 | 5.7 | 853 | 2 T08162 | amylopullulanase (|
| 13 | 101.5 | 5.7 | 705 | 2 T48464 | hypothetical prote |
| 14 | 101.5 | 5.7 | 1396 | 2 F81686 | DNA-directed RNA p |
| 15 | 101 | 5.7 | 394 | 2 F75006 | hypothetical prote |
| 16 | 100 | 5.6 | 336 | 2 H71103 | hypothetical prote |
| 17 | 99.5 | 5.6 | 630 | 2 B64514 | hypothetical prote |
| 18 | 99 | 5.6 | 391 | 2 E64474 | hypothetical prote |
| 19 | 98 | 5.5 | 1430 | 2 D82533 | RNA polymerase bet |
| 20 | 97.5 | 5.5 | 1396 | 2 G71529 | DNA-directed RNA p |
| 21 | 97.5 | 5.5 | 1473 | 2 T38791 | probable ferredoxi |
| 22 | 97 | 5.5 | 480 | 2 E72682 | hypothetical prote |
| 23 | 97 | 5.5 | 924 | 1 JC1269 | DNA-directed DNA p |
| 24 | 96.5 | 5.4 | 819 | 2 E70105 | p15 protein homol |
| 25 | 96.5 | 5.4 | 1367 | 2 C72360 | DNA polymerase III |
| 26 | 96 | 5.4 | 1210 | 2 E64979 | hypothetical 138.1 |
| 27 | 96 | 5.4 | 1242 | 2 S51246 | probable DNA repai |
| 28 | 95.5 | 5.4 | 588 | 2 T37542 | conserved hypotet |
| 29 | 95.5 | 5.4 | 684 | 2 T47694 | probable serine/th |

| | | | | | |
|----|------|-----|------|----------|--------------------|
| 30 | 95.5 | 5.4 | 1285 | 2 B72420 | hypothetical prote |
| 31 | 95 | 5.3 | 445 | 2 C83314 | probable two-compo |
| 32 | 94.5 | 5.3 | 317 | 2 F71139 | hypothetical prote |
| 33 | 94 | 5.3 | 471 | 2 E86192 | hypothetical prote |
| 34 | 93.5 | 5.3 | 235 | 2 T47981 | homeobox-leucine z |
| 35 | 93.5 | 5.3 | 641 | 2 B71116 | probable helicase |
| 36 | 93 | 5.2 | 473 | 1 JC4785 | UTP--glucose-1-pho |
| 37 | 93 | 5.2 | 497 | 2 T29791 | hypothetical prote |
| 38 | 93 | 5.2 | 681 | 2 S73550 | DNA polymerase III |
| 39 | 93 | 5.2 | 923 | 2 G83826 | hypothetical prote |
| 40 | 93 | 5.2 | 948 | 2 G83264 | hypothetical prote |
| 41 | 92.5 | 5.2 | 1209 | 2 H85839 | probable regulator |
| 42 | 92.5 | 5.2 | 4589 | 2 T14914 | dynein beta heavy |
| 43 | 92 | 5.2 | 571 | 2 F84956 | acetylactate synth |
| 44 | 91.5 | 5.1 | 520 | 2 C81273 | hypothetical prote |
| 45 | 91.5 | 5.1 | 529 | 2 S76831 | hypothetical prote |

ALIGNMENTS

RESULT 1

E71144

hypothetical protein PH0368 - Pyrococcus horikoshii

C;Species: Pyrococcus horikoshii

C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 08-Sep-2000

C;Accession: E71144

R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sa

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Og

DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophil

A;Reference number: A71000; MUID:98344137

A;Accession: E71144

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-364 <KAW>

A;Cross-references: GB:AP000002; NID:g3236129; PIDN:BAA29442.1; PID:d1030385; PID:g323

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBa

C;Genetics:

A;Gene: PH0368

C;Superfamily: Pyrococcus horikoshii hypothetical protein PH0368

Query Match 69.1%; Score 1229; DB 2; Length 364;
Best Local Similarity 72.2%; Pred. No. 4.5e-86;
Matches 262; Conservative 47; Mismatches 36; Indels 18; Gaps 18;

Qy 1 LRALVFHGNLQYAEIPKSE-PKVIEKAYIPVIETLIKEE-PFGLNITGYTLKFLPKDII- 57

Db 1 MRALIFHGNLQYAEIPKHEISKVIEKSYFPTISELIKREIPGLNITGYSLQFLPQELIH.60

Qy 58 LVKGGIASDLIEIIGTSYT-AILPLLRSRVEAQVQDR-VKEELFEVSPKGFWLPDLA- 114

Db 61 LIKEGIESELIEILGTSYTHAILPLLTLRSIEAQIKRDREIKKEIFEVSPGGFWLPDLAY 120

Qy 115 DPTIIPAILKDNQYEYLFAD-EAMLFSAHLNSAIKPIKPL-PHLIKAQREKFRFYISYLL- 171

Db 121 DPTIIPAILRDNEYEYLFADGEAMLFSAHLNSAIKSIKPLPYLILKAQRGEGFYLYNLLG 180

Qy 172 LRELKRAIKLVFEGKVTLK-VKDIEAVPVVAVNTAVML-IGRLPLMNPKKVASWIEDKN 229

Db 181 LRELKKAINTLTGGKVTLEAVKDIEAIPVWVSINIAIMLGAGRPPLMSPKRVANWIKGD 240

Qy 230 -ILLYGTDIEFIGYRDIAG-RMSVEGLLEVIDELNSELG-PSELKHSGRELYLRTSSWA- 285

Db 241 EILLYGTDIEFLGYRSIAGHKIAISGLMEVLNSELGELCLPRDIRHNGRRRLYLRTSSWAP 300

Qy 286 DKSLRIWREDEGNARLNML-YNMRGELAFLAENSDARGW-PLPERRLDAFRAIYNDRWG- 342

Db 301 DKSLRIWKEDEGNARLNMLTYCMDEGFAFLAENSDARGWEPLPERRLDAFKAIFYWRNE 360

Qy 343 NGE 345

Db 361 NGK 363

RESULT 2
ALDYAT

amylase A (EC 3.2.1.-) - Dictyoglomus thermophilum
C:Species: Dictyoglomus thermophilum
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 18-Jun-1999
C:Accession: S00628; A34969
R:Fukusumi, S.; Kamizono, A.; Horinouchi, S.; Beppu, T.
Eur. J. Biochem. 174, 15-21, 1988
A:Title: Cloning and nucleotide sequence of a heat-stable amylase gene from an anaerobic
A:Reference number: S00628; MUID:88225097
A:Accession: S00628
A:Molecule type: DNA
A:Residues: 1-686 <FUK>
A:Cross-references: EMBL:X07896; NID:g2688; PIDN:CAA30735.1; PID:g2689
A:Accession: A34969
A:Molecule type: protein
A:Residues: 2-13 <FUK2>
C:Genetics:
A:Gene: amyA
C:Superfamily: Dictyoglomus thermophilum amylase A
C:Keywords: glycosidase; hydrolase
F:2-686/Product: amylase A #status experimental <MAT>

Query Match 8.5%; Score 151; DB 1; Length 686;
Best Local Similarity 21.6%; Pred. No. 0.0011;
Matches 91; Conservative 63; Mismatches 111; Indels 156; Gaps 24;

QY 8 GNLOYAEIPKSEPKVIEKA----YIPVIETLKEEPPFGLNI--TGTYLKFLPKD----- 55
Db 18 GNFD-----VIERAYEMSYKPLINFFPKHPDPINVHFSGFLLLWLEKNHPEYFE 68

QY 56 ---IILVKGGIASDLIEIIGTSYTAILPLPLSRVEAQVQRDRVKE-----ELFEV 103
Db 69 KLKIMAERGQIEF-----VSGGFYEPILPIIP-----DKDKVQKIKLNKYIYDKFGQ 116

QY 104 SPKGFWLPE-LADPIIPAILKDNNGEYELFADEAMLFSAHLNSAIKPIKPLHLIKAQREK 162
Db 117 TPKGMWLAERVWEPHLVKYIAEAGIEYVVVDDAHFFSVGL-----K 157

QY 163 RERYISYLLRLRLKAIKL-VFEGKVTLLKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKV 221
Db 158 EEDLFGYYLME--QYKLAVF--PISMKLYL-----IPFADPEET 195

QY 222 ASWI-----EDKN--ILLYGTDIEFIGYRDIAGRMSVEGLLE-----VIDELNS 263
Db 196 ITYLDKFASEDKSKIALLEDDGKFLWPDYRTYVYEEGWLETFVSKIKENFLVTPVNL 255

QY 264 ELCPSSELKHSGRRELYLRTSSW-----ADKSLR----- 290
Db 256 YTYMQRVKPKGR-IYLPASYREMMEWVLPPEAQKELEELVEKLKTENLWDKFSYVKG 314

QY 291 IWRE-----DEGN-ARLNLMLYMRGELAFLAENSDARGWPLPERRL-DAFRAIYND--WR 341
Db 315 FWRNFLAKYDESNHMOKKMLYVWK-----KVQDSPNEEVKEKAMEEVFQGGANDAYWH 367

QY 342 G 342
Db 368 G 368

RESULT 3
G71241

probable alpha-amylase - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: G71241
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hainaka, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil
A:Reference number: A71000; MUID:98344137
A:Accession: G71241
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-633 <KAW>
A:Cross-references: GB:AP000001; NID:g3236128; PIDN:BAA29262.1; PID:g3256579
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBa
C:Genetics:
A:Gene: PH0193
C:Superfamily: Dictyoglomus thermophilum amylase A

Query Match 7.1%; Score 126.5; DB 2; Length 633;
Best Local Similarity 22.3%; Pred. No. 0.075;
Matches 87; Conservative 59; Mismatches 121; Indels 123; Gaps 24;

QY 21 KVIKAYIPVIETLIKEP---FGLNITGYTLKFL----PKDIILVKGGIASDLIE-IIIG 72
Db 25 RAYEKAYRPFLETL-EEYPMNMKVAVHISGLVLEWLENRPEYIDLKSLIKKGQVELVVA 83

QY 73 TSYTAILPLLP-LSRVEAQVQRDRVKEELFEVSPKGFWLPE-LADPIIPAILKDNNGEYEL 130
Db 84 GFYEPILVAIPEEDRVE-QIKLSKGWARKMGYEARGLWLTERTVWEPELVKTLREAGIEYV 142

QY 131 FADEAMLFSAHLNS-----AIKPIKPLPHLIKAQREKRRERYISYLLRE 174
Db 143 ILDDYHMSAGLSKEELFWPYTENGGEAIVVFPI-----DEKRLYL--IPFRP 189

QY 175 LRKAIKLVF-----EGKVTLLKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWI--- 225
Db 190 VNETLEYLHSLADEDESKVAVFHDDGKFGAWPGTHLVYERG-----WLKEF 237

QY 226 -----EDK-NILLYGTDIEFIG-----YRDIAG--RMSVEGL-----LEV 257
Db 238 FDRISDDKINMLYS---EYLSKFRPKGLVYLPPIASYFEMSEWSLPAQAKLFFEFIKK 294

QY 258 IDELNSLCPSELKHSGRRELYLRTSSWADKSLRIWREDEGN---ARLNLMLYMRGELAF 314
Db 295 LKELN-----LFEKYRIFVRGGIWKN---FLYKYPEGNYMHKRLMLSKL-----L 337

QY 315 AENSARGWPLPERRLDAFRAIYND--WRG 342
Db 338 RNN-----PTARIFVLRAQCNDAYWHG 359

RESULT 4
E75206

alpha-amylase (or 4-alpha-glucanotransferase) PAB0118 - Pyrococcus abyssi (strain Ors
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: E75206
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: E75206
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-655 <KAW>
A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49100.1; PID:g545
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: amyA; PAB0118
C:Superfamily: Dictyoglomus thermophilum amylase A

Query Match 6.8%; Score 120.5; DB 2; Length 655;
Best Local Similarity 20.1%; Pred. No. 0.22;
Matches 79; Conservative 62; Mismatches 117; Indels 135; Gaps 20;

Db 3641 -----NWHANSLP 3648

RESULT 7

A75207

amylopullulanase PAB0122 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: A75207

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure

A:Reference number: A75001

A:Accession: A75207

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1362 <KAW>

A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49104.1; PID:el51499

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: apu; PAB0122

Query Match 6.2%; Score 110; DB 2; Length 1362;

Best Local Similarity 21.9%; Pred. No. 3.9;

Matches 93; Conservative 61; Mismatches 150; Indels 120; Gaps 22;

QY 11 QYAEIPKSEPKV-----JEKAYIPV----- 30

Db 174 KYSNLPLEEQKIAVTSEFTQDYIDLAVLFNLAWIDYIMNTPELKALYDKVDTGGYTR 233

QY 31 --IETLIKEEPFGLNITGYTLKFLPKDIIILVKGGIASDLIEIIGTSYTAILPLPL--- 84

Db 234 KDVETVLKHQMWLLN---HTFEEHEKINLLGNG---NVEVTVPYTH--PIGPILNDF 284

QY 85 ---SRVEAQVQR--DRVKEEL--FEVSPKGFWLPE--LADPIIPAILKDNNGYEYLFAD 135

Db 285 GWYEDFDAQVKKANELYKEYLGAGKVTTPKCGWAESAINDKTL-EILAENGWKWMTDQL 343

QY 136 MLFSAHLNSAI-----KPIKPLPHLIKAQREKFRYISY-----LLLREL 175

Db 344 VLEKLGVPKTIESYYKPPVQAQFGDKKIYLFPRNHDLSDRVGFYAGMNQYDAVKNFVEEL 403

QY 176 RKAIKLVFEGKVTLVKVDIEAVPVWAVNTAVMLIGRLPLMNPKKVASWIEDKNILLYGT 235

Db 404 LKIQKQNYDGLS-VYVITLDGENPW---EHVPFDGKLFLEELYRQLEELQKGLIRTVT 458

QY 236 DIEFIG-YRDIAGRM-----SVEGLL--EVIDELNSELCPSELKHSGRELY 278

Db 459 PSEYIEMFGDKANKLTPKMMKRLDFTTEDNVALLKAKTLGELYDMVGVT-----EMQ 512

QY 279 LRTSSWADKSLRIW-REDEGNARLNMLYNNRGEAFLAENS-ARGWPLPERRLDARAI 336

Db 513 WPRESSWIDGTLSWTIGEPQENIAWYWLILAR---KALFENKDNVKNKAYEYL--FRAE 567

QY 337 YNDW 340

Db 568 GSDW 571

RESULT 8

H82815

fimbrial assembly membrane protein XF0369 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: H82815

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing

Nature 406, 151-157, 2000

A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717

A>Note: for a complete list of authors see reference number A59328 below

A:Accession: H82815

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-392 <SIM>

A:Cross-references: GB:AE003888; GB:AE003849; NID:g9105187; PIDN:AAF83179.1; GSPDB:GN

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R. Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C. F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; de Silv M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L. A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF0369

Query Match 6.0%; Score 106.5; DB 2; Length 392;

Best Local Similarity 24.2%; Pred. No. 1.3;

Matches 80; Conservative 34; Mismatches 97; Indels 119; Gaps 16;

QY 27 YIPVETLIKEEPF-GLNITGYTLKFLPKDIIILVKGGIASDLIEIIGTSYTAILPLPLS 85

Db 27 YVPVLKVPVVRGVFVGL-----FSKKQSVLVGVDSITAVK-----LLQLS 67

QY 86 RVEAQVQRDRVKKEELFEVSP-----KGFWLPELADPIIPAILKDNNGYEYLFAD 137

Db 68 R-----SGNRFKVEHYAVEPLPLNAVAERGIVEQVGEAIRRAVRSRGTAKAKFAA 122

QY 138 FSAHLNSAIKPIKPLP-----HLIKAQRE-KRFRYISY----- 169

Db 123 GSAVITKLI-----PMPVGLDEQDLAEQIEATNPIPIEEVSLDFEVLGPVNNTEMV 178

QY 170 -LLLR-----ELRK-AIKLVFEGKVTLVKVDIEAVPVWAVNTAVMLIGRLPLMNP 220

Db 179 OVLLAASRSENVLRQSALEL---GGLTAKVIDVEA-----LAVENAFSLIAQELSVGSNA 231

QY 221 VASWIEDKNILLYGTDIEFIGYRDIAGRMSVEGLLEVIDELNSELCPSELKHSGRELYLR 280

Db 232 LVALI-----DIGATMSTLNL-----HSGRSLYTR 257

QY 281 TSSWADKSLRIWREDEGNARLNMLYNNRGE 310

Db 258 EQLFGGKQL-----TDEVHRYGMYEAGQ 283

RESULT 9

B64501

alpha-amylase (EC 3.2.1.1) - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: B64501

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak ; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese A>Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc A:Reference number: A64300; MUID:96337999

A:Accession: B64501

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-467 <BUL>

A:Cross-references: GB:U67601; GB:L77117; NID:g2826439; PIDN:AAB99631.1; PID:g1592212

C:Genetics:

A:Map position: FOR1586396-1587799

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

Db 266 ALTSPKFAFKARNWELKGTESIDVTDVAGSNIRIDSRGPEVMRVVPRNLNEDINEEWISDK 325
 QY 229 NILLY-----GTDIEF--IGYRDIAGRMSVEGLEVIDELNSELCPSEL-- 270
 Db 326 TRFFYDGLKRQRLNDPMIRGADGRFOQVSWRD-----ALAIVAEVMHQIKPEEIVG 376
 QY 271 ---KHSGRELYLRTSSWADK--SLRIWREDEG---NARLNMLNMRGELAFLAENSDA 320
 Db 377 VAGKLSDAESMMALKDLLNKMGSNNIFCEGNGMHPNADLRSGYIMNTSISGL-EKADA 433

RESULT 12

T08162
 C:Species: Pyrococcus furiosus
 C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 08-Oct-1999
 C:Accession: T08162
 R: Dong, G.; Vieille, C.; Zeikus, J.G.
 Appl. Environ. Microbiol. 63, 3577-3584, 1997
 A: Title: Cloning, sequencing, and expression of the gene encoding amylopullulanase from
 A: Reference number: Z16389; MUID: 97438521
 A: Status: preliminary
 A: Molecule type: DNA
 A: Molecule type: translated from GB/EMBL/DDBJ
 A: Residues: 1-853 <DON>
 A: Cross-references: EMBL: AF016588; NID: g2435436; PIDN: AAB71229.1; PID: g2435438
 C: Genetics: apu
 C: Function:
 A: Description: hydrolyzes alpha-1,6 and alpha-1,4 linkages in starch-related polysacchar
 C: Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 5.7%; Score 102; DB 2; Length 853;
 Best Local Similarity 21.2%; Pred. No. 8.3;
 Matches 90; Conservative 60; Mismatches 155; Indels 120; Gaps 22;

QY 11 QYAEIPKSEPKV-----IEKAYI-----PVIETLIKEEPFGLNIT 45
 Db 172 KYANLPLEEQKVAVTNEFTQDYIDLAVLFNLAWIDYNIISTPELKALYDKVDEG---- 227
 QY 46 GYTLKFLPK-----DIILVKGGIASDLIEIGTSYTAILPLPL-- 84
 Db 228 GYTREDLKTLYHQMWLLNNTFKEHEKINLLGNGN-----VEVTVPYAH--PIGPILN 280
 QY 85 -----SRVEAQVQDRVKEELFE-----VSPKGFWLPE--LADPIIPAILKDNNGEYL 130
 Db 281 DFGWSEDFDAHVKK---AHELYKKYLGCGVATPRGGWAAESALNDKTL-EILAENGWQWV 336
 QY 131 FADEAML-----FSAHLNSAIKPIKPLPHLIKAQREKFRFYISYL----- 170
 Db 337 MTQDMVLRMGIPYSIENYRPPWAFENG--KKIYLFPRNHDLSDRVGFYSGMNQYEA 394
 QY 171 --LLRELKRAIKLVFEGKVTILKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVAVSWIEDK 228
 Db 395 EDFINELKIQKYNIDGSL-VYVITLDGENPW---EHYPYDGKLFLETLYKRLSELQEA 449
 QY 229 NILLYGTDIEFIG-YRDIAGRMSVEGLEVIDELNSELCPSEL----- 278
 Db 450 GLIRTLTPTEYIQLYGDKANKLTPQ-MMERLDTTEERVEALKVANSGLGELYDLAGVTEE 508
 QY 279 --LRTSSWADKSLRIW-REDEGNARLNMLNMRGELAFLAENSDARGWPLPERLDAFRA 335
 Db 509 MQWPESWIDGTLSWTWIGEPQENYAWYWLYLARRTLNENKDKMDSASWEKAYEYL--LRA 566
 QY 336 IYNDW 340
 Db 567 EASDW 571

RESULT 13

T48464
 hypothetical protein T1E3.40 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48464
 R: Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.;
 ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, April 2000
 A: Reference number: Z24491
 A: Accession: T48464
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-705 <BEV>
 A: Cross-references: EMBL: AL162972
 A: Experimental source: cultivar Columbia; BAC clone T1E3
 C: Genetics:
 A: Map position: 5
 A: Introns: 59/1; 286/1; 350/1; 493/3
 A: Note: T1E3.40

Query Match 5.7%; Score 101.5; DB 2; Length 705;
 Best Local Similarity 21.2%; Pred. No. 6.9;
 Matches 81; Conservative 59; Mismatches 141; Indels 101; Gaps 20;

QY 1 LRALVFHGNLQYAE-IPKSEPKVIE-----KAYIPVIETLIKEEPFGLNITGYTLKFLPKD 55
 Db 143 LTVVAVSGNMEIAEALVAKNPKLLEIPGINGQIPV---VAVENTQMEMARYLYTRTPVQ 199
 QY 56 IILVKGGIASDLI-----EIIGTSYTAILPLPLSRVEAQ-----VQDRVKEELFEVS 104
 Db 200 VLLAEDGYHGTLFLFLNAIFYRMLGKGLGIQATHIFGGFDLYLFFFIQLD-IALDLFNMS 258
 QY 105 PKGFWLPELADPIIPAILKDNNGEYVLFADAEAMLFSAHLNSAIK-PIKPLP-HLIKAQREK 162
 Db 259 RRLAVTKHLQIESIPII-----VLASKPDLPF---GIOVKLPTLPKPSHANKDHKSK 307
 QY 163 RFRYS-----YLLRLRELKRAIKLVFEGKVTILKVKDIEAVPVVAVNTAVMLIGRLPLM 216
 Db 308 FRIHKVYKKSIIYIPLKKVKRSFDLPD---TLMRKLLKGLSKWTGIDE----VYRLKVM 360
 QY 217 N-PKKVASWIEDKNILL-----YGTDFIEFIGYRDIAGRMSVEGLE 256
 Db 361 HLQAKKLLLGISEETLTGLKERSETVDEALFEAVRYG-----NVDFLVE 405
 QY 257 VIDELNSELCPSELKHSGRELYLRTSSWADKSLRIWREDEGNARLNMLNLM--RGELAF 314
 Db 406 MTRN-NSELLWSTRTSSSTLFLLAVERQEKV-----FNLLYGLDDRKYLLA 453
 QY 315 AENSAR-----CWPLPERRL 330
 Db 454 DKDSGNGVVLHLAGFPSPPSKL 475

RESULT 14

F81686
 DNA-directed RNA polymerase, beta' chain TC0588 [imported] - Chlamydia muridarum (str
 C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
 C:Accession: F81686
 R: Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
 , C.; Dodson, R.; Winn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
 Nucleic Acids Res. 28, 1397-1406, 2000
 A: Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
 A: Reference number: A81500; MUID: 20150255
 A: Accession: F81686
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-1396 <TET>
 A: Cross-references: GB: AE002328; GB: AE002160; NID: g7190627; PIDN: AAF39420.1; PID: g719
 A: Experimental source: strain Nigg (MoPn)
 C: Genetics:
 A: Gene: TC0588
 C: Superfamily: Escherichia coli DNA-directed RNA polymerase beta' chain

| Result No. | Score | Query % | | Length | DB | ID | Description |
|------------|-------|---------|--|--------|----|-------------|---------------------|
| | | Match | | | | | |
| 1 | 151 | 8.5 | | 685 | 1 | AMY1_DICTH | P09961 dictyoglomu |
| 2 | 126.5 | 7.1 | | 633 | 1 | AMYA_PYRHO | O57932 pyrococcus |
| 3 | 120.5 | 6.8 | | 655 | 1 | AMYA_PYRAB | Q9v298 pyrococcus |
| 4 | 116.5 | 6.6 | | 648 | 1 | AMYA_PYRUF | P49067 pyrococcus |
| 5 | 111 | 6.2 | | 4725 | 1 | DYHC_DICDI | P34036 dictyosteli |
| 6 | 106.5 | 6.0 | | 467 | 1 | AMYA_METJA | Q59006 methanococ |
| 7 | 104.5 | 5.9 | | 1250 | 1 | BXE_CLOBU | P30995 clostridium |
| 8 | 102 | 5.7 | | 738 | 1 | NUAM_SOLTU | Q43644 solanum tub |
| 9 | 101.5 | 5.7 | | 1396 | 1 | RPOC_CHLMU | Q9pk79 chlamydia m |
| 10 | 100.5 | 5.7 | | 653 | 1 | MALQ_PYRKO | O32450 pyrococcus |
| 11 | 99.5 | 5.6 | | 630 | 1 | YZ35_METJA | Q60291 methanococc |
| 12 | 97.5 | 5.5 | | 659 | 1 | MALQ_THELI | O32462 thermococcu |
| 13 | 97.5 | 5.5 | | 1396 | 1 | RPOC_CHLTR | O84316 chlamydia t |
| 14 | 97 | 5.5 | | 924 | 1 | DPOL_BPSP1 | P30314 bacterioph |
| 15 | 96.5 | 5.4 | | 1367 | 1 | DPO3_THEMEA | Q9zhf6 thermotoga |
| 16 | 96 | 5.4 | | 1210 | 1 | YEHI_ECOLI | P33346 escherichia |
| 17 | 96 | 5.4 | | 1242 | 1 | MSH6_YEAST | Q03834 saccharomyc |
| 18 | 95.5 | 5.4 | | 471 | 1 | UDPG_PYRPY | O64459 pyrus pyrifi |
| 19 | 94 | 5.3 | | 467 | 1 | UDPG_MUSAC | Q9sdx3 musa acumin |
| 20 | 93 | 5.2 | | 473 | 1 | UDPG_HORVU | Q43772 hordeum vul |
| 21 | 93 | 5.2 | | 681 | 1 | DP3X_MYCPN | P75177 mycoplasma |
| 22 | 92 | 5.2 | | 571 | 1 | ILVI_BUCAI | P57321 buchnera ap |
| 23 | 91 | 5.1 | | 598 | 1 | Y288_THEMEA | Q9wyc4 thermotoga |
| 24 | 90.5 | 5.1 | | 781 | 1 | YG5B_YEAST | P53316 saccharomyc |
| 25 | 90 | 5.1 | | 470 | 1 | PPOX_BACSU | P32397 bacillus su |
| 26 | 89.5 | 5.0 | | 387 | 1 | HEMZ_SYNY3 | P54225 synecocyst |
| 27 | 89.5 | 5.0 | | 446 | 1 | SR54_BACSU | P37105 bacillus su |
| 28 | 89.5 | 5.0 | | 1078 | 1 | RPOB_SINAL | P46818 sinapis alb |
| 29 | 88.5 | 5.0 | | 447 | 1 | SR54_MYCMI | Q01442 mycoplasma |
| 30 | 88 | 4.9 | | 257 | 1 | VBR1_TMOV | Q06661 tomato mott |
| 31 | 88 | 4.9 | | 572 | 1 | PT1_STAAN | P51183 staphylococ |
| 32 | 88 | 4.9 | | 592 | 1 | STB3_HUMAN | O00186 homo sapien |
| 33 | 87.5 | 4.9 | | 476 | 1 | UDPG_SOLTU | P19595 solanum tub |

Query Match 7.1%; Score 126.5; DB 1; Length 633;
Best Local Similarity 22.3%; Pred. No. 0.026;
Matches 87; Conservative 59; Mismatches 121; Indels 123; Gaps 24;

QY 56 ---IILVKGGIASDLIEITGTSYTAILPLPLSRVQAQVDRVKE-----ELFEV 103
Db 68 KIKMAERGQIEF---VSGGFYEPILPIIP-----DKDKVQKIKLKYIYDKFGQ 115

QY 104 SPKGFWLPE-LADPIIPAILKNDNGYEYLFADAEAMLFSAHLNSAIKPIKPLPHLIKAQREK 162
Db 116 TPKGMWLAERVWEPHLVYIAEAGIEYVVVDAAHFFSVGL-----K 156

QY 163 RFRYISYLLRELKAIKL-VFEGKVTLKVKDIEAVPVVAVNTAVMLTGRPLPLMNPVKV 221
Db 157 EEDLFGYILMEE--QGYKLAVF--PISMRLRYL-----IPFADPEET 194

QY 222 ASWI-----EDKN--ILLYGTDIEFIGYRDIAGRMSVEGLLE-----VIDELNS 263
Db 195 ITYLDKFASEDKSKIALDFDGEKFGLPDPTRTVYEEGWLETFVSKIKENFLLVTPVNL 254

QY 264 ELCPSSELKHSRELYLRTSSW-----ADKSLR----- 290
Db 255 YTYMQRVKPKGR-IYLPATASYREMMEWVLFPEAQKELEELVEKLKTENLWDKFSYVKG 313

QY 291 IWRE-----DEGN-ARLNMLYNMRGELAFLAENSDARGWPLPERRL-DAFRAIYND--WR 341
Db 314 FWRNFLAKYDESNHMQKMLYVWK-----KVQDSPNEEVKEMEEVFQGOANDAYWH 366

QY 342 G 342
Db 367 G 367

RESULT 2
AMYA_PYRHO
ID AMYA_PYRHO STANDARD; PRT; 633 AA.
AC 057932;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ALPHA-AMYLASE (EC 3.2.1.1).
GN AMYA OR PHO193.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudo H. Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -!- PATHWAY: POLYSACCHARIDE DEGRADATION.
CC -!- SIMILARITY: BELONGS TO FAMILY 57 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP000001; BAA29262.1; -
DR Hydrolase; Glycosidase; Carbohydrate metabolism; Complete proteome.
KW
SQ SEQUENCE 633 AA; 75017 MW; 9DCBCBB2FE191501 CRC64;

QY 21 KVIEKAYIPVETLIIKEEP---FGLNITGYTLKFL-----PKDIILVKGGIASDLIE-IIG 72
Db 25 RAVEKAYRPFLETL-EEYPNMKVAVHISGLVLEWERNRPEYIDLKSLIKKGQVELVVA 83

QY 73 TSYTAILPLLP-LSRVEAQVQDRVKEELFEVSPKGFWLPE-LADPIIPAILKNDNGYEYL 130
Db 84 GFYEPILVAIPEEDRVE-QIKLSKGWARKMGVEARGLWLTERTVWEPELVKTLREAGIEYV 142

QY 131 FADEAMLFSAHLNS-----AIKPIKPLPHLIKAQREKRRFRVISYLLRE 174
Db 143 ILDDYHFMASGLSKEELFWPYTYTENGGEAIVVFP-----DEKLRYL--IPFRP 189

QY 175 LRKAIAKLVF-----EGKVTLKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWI--- 225
Db 190 VNETLEYLHSLADESKVAVFHDDGKFGAWPGTHLVYER-----WLKEF 237

QY 226 -----EDK-NILLYGTDIEFIG-----YRDIAG--RMSVEGL-----LEV 257
Db 238 FDRISDDKINLMYS---EYLSKFRPKGLVYLPPIASYFEMSEWSLPARQAKLFFFIKK 294

QY 258 IDELNSELCPSELKHSRELYLRTSSWADKSLRIWREDEGN---ARLNMLYNMRGELAF 314
Db 295 LKELN-----LFEKYRIFVRGGIWKN---FLYKYPEGNYMHKRLMLSKL-----L 337

QY 315 AENSARGWPLPERRLDAFRAIYND--WRG 342
Db 338 RNN-----PTARIFVLRAQCNDAYWHG 359

RESULT 3
AMYA_PYRAB
ID AMYA_PYRAB STANDARD; PRT; 655 AA.
AC Q9V298;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ALPHA-AMYLASE (EC 3.2.1.1).
GN AMYA OR PAB0118.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -!- PATHWAY: POLYSACCHARIDE DEGRADATION.
CC -!- SIMILARITY: BELONGS TO FAMILY 57 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ248283; CAB49100.1; -
DR Hydrolase; Glycosidase; Carbohydrate metabolism; Complete proteome.
KW
SQ SEQUENCE 655 AA; 77296 MW; 7F6F920B1A00EECE CRC64;

Query Match 6.8%; Score 120.5; DB 1; Length 655;
Best Local Similarity 20.1%; Pred. No. 0.079;
Matches 79; Conservative 62; Mismatches 117; Indels 135; Gaps 20;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 14:40:16 ; Search time 28.84 seconds
(without alignments)
1754.862 Million cell updates/sec

Title: US-09-407-806A-4
Perfect score: 1778
Sequence: 1 LRALVFHGNLQYAEIPKSEP.....ERRLDAFRAIYNDRNGNEP 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_invertebrate:
14: sp_unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|-----------|----------------------|
| 1 | 1268 | 71.3 | 364 | 1 Q9HHB5 | Q9hhb5 pyrococcus |
| 2 | 1229 | 69.1 | 364 | 1 O58106 | O58106 pyrococcus |
| 3 | 110 | 6.2 | 1362 | 1 Q9V294 | Q9v294 pyrococcus |
| 4 | 106.5 | 6.0 | 392 | 2 Q9PGD3 | Q9pgd3 xylella fas |
| 5 | 106 | 6.0 | 458 | 10 Q9MBD0 | Q9mbd0 pyrus pyrif |
| 6 | 104.5 | 5.9 | 1255 | 2 Q9FAR6 | Q9far6 clostridium |
| 7 | 102 | 5.7 | 853 | 1 O30772 | O30772 pyrococcus |
| 8 | 101.5 | 5.7 | 705 | 10 Q9LZ29 | Q9l229 arabisidopsis |
| 9 | 101 | 5.7 | 394 | 1 Q9UYC0 | Q9uyc0 pyrococcus |
| 10 | 100.5 | 5.7 | 653 | 1 O32450 | O32450 pyrococcus |
| 11 | 100.5 | 5.7 | 690 | 2 Q9PKB1 | Q9pkb1 chlamydia m |
| 12 | 100 | 5.6 | 336 | 1 O58818 | O58818 pyrococcus |
| 13 | 99 | 5.6 | 391 | 1 Q58793 | Q58793 methanococc |
| 14 | 99 | 5.6 | 608 | 10 Q9AVF4 | Q9avf4 amaranthus |
| 15 | 99 | 5.6 | 742 | 5 O17474 | O17474 hydra oliga |
| 16 | 98.5 | 5.5 | 1518 | 13 Q9IAR8 | Q9iar8 salmo salar |
| 17 | 98 | 5.5 | 1430 | 2 Q9PA87 | Q9pa87 xylella fas |
| 18 | 97.5 | 5.5 | 564 | 2 Q99WX0 | Q99wx0 staphylococ |
| 19 | 97.5 | 5.5 | 659 | 1 O32462 | O32462 thermococcu |

| | | | | | | |
|----|------|-----|------|----|--------|---------------------|
| 20 | 97.5 | 5.5 | 1473 | 3 | O14167 | O14167 schizosacch |
| 21 | 97.5 | 5.5 | 1503 | 13 | O73677 | O73677 fundulus he |
| 22 | 97 | 5.5 | 480 | 1 | Q9YDN6 | Q9ydn6 aeropyrum p |
| 23 | 97 | 5.5 | 1511 | 13 | Q9I8E2 | Q9i8e2 fugu rubrip |
| 24 | 96.5 | 5.4 | 819 | 2 | O51074 | O51074 borrelia bu |
| 25 | 95.5 | 5.4 | 588 | 3 | O13693 | O13693 schizosacch |
| 26 | 95.5 | 5.4 | 684 | 10 | Q9M2S4 | Q9m2s4 arabisidopsi |
| 27 | 95.5 | 5.4 | 1285 | 2 | Q9WXU3 | Q9wxu3 thermotoga |
| 28 | 95 | 5.3 | 445 | 2 | Q9I0I2 | Q9i0i2 pseudomonas |
| 29 | 95 | 5.3 | 1337 | 1 | Q9Y8I8 | Q9y8i8 thermococcu |
| 30 | 94.5 | 5.3 | 317 | 1 | O58067 | O58067 pyrococcus |
| 31 | 94 | 5.3 | 471 | 10 | Q9MA46 | Q9ma46 arabisidopsi |
| 32 | 94 | 5.3 | 1251 | 2 | Q9K395 | Q9k395 clostridium |
| 33 | 93.5 | 5.3 | 235 | 10 | Q9M276 | Q9m276 arabisidopsi |
| 34 | 93.5 | 5.3 | 641 | 1 | O58428 | O58428 pyrococcus |
| 35 | 93 | 5.2 | 497 | 5 | Q27420 | Q27420 caenorhabdi |
| 36 | 93 | 5.2 | 923 | 2 | Q9KD04 | Q9kd04 bacillus ha |
| 37 | 93 | 5.2 | 948 | 2 | Q9HZE4 | Q9hze4 pseudomonas |
| 38 | 92.5 | 5.2 | 4589 | 5 | O76506 | O76506 tetrahymena |
| 39 | 91.5 | 5.1 | 364 | 2 | O31292 | O31292 buchnera ap |
| 40 | 91.5 | 5.1 | 520 | 2 | Q9PM27 | Q9pm27 campylobact |
| 41 | 91.5 | 5.1 | 529 | 2 | P74630 | P74630 synechocyst |
| 42 | 91.5 | 5.1 | 794 | 2 | Q9X2I2 | Q9x2i2 thermotoga |
| 43 | 91 | 5.1 | 510 | 2 | O26022 | O26022 helicobacte |
| 44 | 91 | 5.1 | 512 | 4 | Q9H5U5 | Q9h5u5 homo sapien |
| 45 | 91 | 5.1 | 825 | 5 | O16769 | O16769 caenorhabdi |

ALIGNMENTS

RESULT 1
Q9HHB5
ID Q9HHB5 PRELIMINARY; PRT; 364 AA.
AC Q9HHB5;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)
DE ALPHA-GALACTOSIDASE.
GN GALA.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 3638;
RA Verhees C.H.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF195244; AAG28455.1; -
SQ SEQUENCE 364 AA; 41545 MW; 3E8E0AF5BDCCF2A5 CRC64;

| | | | | |
|-----------------------|------------------|--|------------|-------------|
| Query Match | 71.3% | Score 1268; | DB 1; | Length 364; |
| Best Local Similarity | 75.2% | Pred. No. 2.2e-81; | | |
| Matches 273; | Conservative 38; | Mismatches 34; | Indels 18; | Gaps 18; |
| Qy | 1 | LRALVFHGNLQYAEIPKSE-PKVIEKAYIPVIETLIKEE-PFGLNITGYTLKFLPKDII- | 57 | |
| Db | 1 | MRALVFHGNLQYAEIPKSEIPKVIEKAYFPTISELIRREIPFPGLNITGYSLSLFKDLIA | 60 | |
| Qy | 58 | LKGGIASDLIEIIGTSYT-AILPLPLSRVEAQVQRDR-VKEELFEVSPKGFWLPDLA- | 114 | |
| Db | 61 | LIKEGIESGLIEILGTSYTHAILPLPLSRVEAQIKRDREKVENILEVSPGFWLPDLAY | 120 | |
| Qy | 115 | DPIIPAILKDNQGYEYLFAD-EAMLSAHLNSAIKPIKPL-PHLIKAQREKFRFYISYLL- | 171 | |
| Db | 121 | DPIIPAILRDNNYEYLFADGEAMLSFNLNSAIKPIKPLYPHLIKAQREGGLVLYNLLG | 180 | |
| Qy | 172 | LRELRAIKLVFEKGKVTLK-VKDIEAVPVWVAVNTAVML-IGRLPLMNPKKVASWIEDKN | 229 | |
| Db | 181 | LRELKKAINLVFEKGKVTLEAVKEIEAIPVWVSINTAVMLGAGRFPLMNPKKVAKWKEKD | 240 | |
| Qy | 230 | -ILLYGTDIEFIGYRDIAG-RMSVEGLEVIDELNSEL-CPSELKHSGRELYLRTSSWA- | 285 | |

Db 241 EILLYGTDFIEFLGYRDIAGYKITISNLEIINELEGEGLPRKIKHSEKKLYLRTSSWAP 300
QY 286 DKSRLRWREDEGNARLNMLYN-MRGELAFLAENSDARGW-PLPERRLDAFRAIYNDWRG- 342
Db 301 DKSRLRWTEDEGNARLNMLTSYMDGELAFLAENSDARGWEPLPERRLDAFAKAIYTHWRSE 360
QY 343 NGE 345
Db 361 NGK 363

RESULT 2
O58106 PRELIMINARY; PRT; 364 AA.
AC O58106;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 41.8 KDA PROTEIN PH0368.
GN PH0368.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL; AP000002; BAA29442.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 364 AA; 41755 MW; 7B4B36AB4A975BAD CRC64;

Query Match 69.1%; Score 1229; DB 1; Length 364;
Best Local Similarity 72.2%; Pred. No. 1.2e-78;
Matches 262; Conservative 47; Mismatches 36; Indels 18; Gaps 18;

QY 1 LRALVFHGNLQYAEIPKSE-PKVIKAYIPVIETLIKEE-PFGLNITGYTLKFLPKDII- 57
Db 1 MRALIFHGNLQYAEIPKSEISKVIEKSYFPTISELIKREIPFGLNITGYSLQFLPQELIH 60
QY 58 LVKGGIASDLIEIIGTSYT-AILPLLPLSRVEAQVQRDR-VKEELFEVSPKGFWLPELA- 114
Db 61 LIKEGIESELIEILGTSYTHAILPLLTLSRIEAIKRDREIKEEIFEVSPGGFWLPCLAY 120
QY 115 DPLIIPAILKDNNGEYELFAD-EAMLFSAHLNSAIKPIKPL-PHLIKAQREKRFYISYLL- 171
Db 121 DPLIIPAILRDNEYEYELFADGEAMLFNSHLNSAIKSIKPLYPYLIKAQRGEGFVLYNLLG 180
QY 172 LRELKAIKLVFEGKVTLK-VKDIEAVPVVAVNTAVML-IGRLPLMNPKKVASWIEDKN 229
Db 181 LRELKAINLTFGGKVTLEAVKDIEAIPVWVSINIAIMLGAGFRPLMSPKRVANWIKGD 240
QY 230 -ILLYGTDIEFIGYRDIAG-RMSVEGLLEVIDELNSELC-PSELKHSGRRELYLRTSSWA- 285
Db 241 EILLYGTDIEFLGYRSIAGHKIAISGLMEVLNGLGELCLPRDIRHNGRRLYLRTSSWAP 300
QY 286 DKSRLRWREDEGNARLNML-YNMRGELAFLAENSDARGW-PLPERRLDAFRAIYNDWRG- 342
Db 301 DKSRLRWKEDEGNARLNMLTYCMDGEFAFLAENSDARGWEPLPERRLDAFAKAIYKYWRNE 360
QY 343 NGE 345
Db 361 NGK 363

RESULT 3
Q9V294
ID Q9V294 PRELIMINARY; PRT; 1362 AA.
AC Q9V294;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE AMYLOPULLULANASE.
GN APU OR PAB0122.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ248283; CAB49104.1; -.
KW Complete proteome.
SQ SEQUENCE 1362 AA; 154578 MW; 277AFAB4E14860D1 CRC64;

Query Match 6.2%; Score 110; DB 1; Length 1362;
Best Local Similarity 21.9%; Pred. No. 40;
Matches 93; Conservative 61; Mismatches 150; Indels 120; Gaps 22;

QY 11 QYAEIPKSEPKV-----IEKAYIPV----- 30
Db 174 KYNLPLEEQKIATVSEFTEQDYIDLAVLFNLAWIDYIMNTPELKALYDKVDTGGYTR 233
QY 31 --IETLIKEEPFGLNITGYTLKFLPKDITLVKGGIASDLIEIIGTSYTAILPLLPL---- 84
Db 234 KDVETVLKHQMWLLN--HTFEEHEKINLLGNG---NVEVTVPYTH--PIGPILNDF 284
QY 85 ---SRVEAQVQR--DRVKEEL--FEVSPKGFWLPE--LADPIIPAILKDNNGEYELFADEA 135
Db 285 GWYEDFDAQVKKANELYKEYLGAGKVTPKGGWAAESALNDKTL-EILAENGKWKWMTDQL 343
QY 136 MLFSAHLNSAI-----KPIKPLPHLIKAQREKRFYISY-----LLLREL 175
Db 344 VLEKLGVPKTIETSYKPPWVAQFGDKKIYLFPRNHDLSDRVGFYAGMNQYDAVKNFVEEL 403
QY 176 RKAIKLVFEGKVTLKVDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWIEDKNILLYGT 235
Db 404 LKIQKQNYDGLS-VYVITLDGENPW----EHPDFDGKLFLEELYRQLEELQKKGLIRTVT 458
QY 236 DIEFIG-YRDIAGRM-----SVEGLL--EVIDELNSELCPSSELKHSGRRELY 278
Db 459 PSEYIEMFGDKANKLTPKMMKRLDFTEDNVNALLKAKTLGELYDMVGVT-----EMQ 512
QY 279 LRTSSWADKSLRIW-REDEGNARLNMLYNMRGELAFLAENSD-ARGWPLPERRLDAFRAI 336
Db 513 WPESWIDGTLSTWIGEPQENIAWYLYLAR---KALFENKDNVKNKAYEYL--FRAE 567
QY 337 YNDW 340
Db 568 GSDW 571

RESULT 4
Q9PGD3
ID Q9PGD3 PRELIMINARY; PRT; 392 AA.
AC Q9PGD3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FIMBRIAL ASSEMBLY MEMBRANE PROTEIN.
GN XF0369.
OS Xylella fastidiosa.

[illegible]

Db 200 VLLAEDGYHGTLFLNAIFYRMLGKGLGFIQATHIFGGFDLYLFFFIQLD-IALDLFNMS 258
QY 105 PKGFWLPADPIIPAILKNDGYEYLFADAMLSAHLNSAIK-PIKPLP-HLIKAQREK 162
Db 259 RRLAVTKHLQIESIPII-----VLASKPDLPF---GIQVKLPTLPKPSHANKDHKSK 307
QY 163 RFRYIS-----YLLRLRELKAIKLVFEGKVTLKVKDIEAVPVVAVNTAVMLIGRLPLM 216
Db 308 FFRHKVYKKSIIYIPLKKVKRSFDLPD---TLMRKLKGLSKWTGIDE---VYRLKVM 360
QY 217 N--PKKVASWIEDKNILL-----YGTDIEFIGYRDIAGRMSVEGLLE 256
Db 361 HLQAKLLLGISEETLTGLKERSETVDEALLFAVRG-----NVDPLVE 405
QY 257 VIDELNSELCPSELKHSGRRELYLRTSSWADKSLRIWREDEGNARLNMLYNM--RGELAF 314
Db 406 MIRN-NSELLNWSTRSSSSTFLLAVERQEKV-----FNLLYGLDDRKYLLLA 453
QY 315 AENSAR-----GWPLPERRL 330
Db 454 DKDSGNGVHLHLAGFPSPPSKL 475

RESULT 9
Q9UYC0 PRELIMINARY; PRT; 394 AA.
AC Q9UYC0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 44.0 KDA PROTEIN.
GN PAB1309.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ248288; CAB50492.1;
DR InterPro; IPR001696; Na_channel.
DR PRINTS; PR00170; NACHANNEL.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 394 AA; 43951 MW; 11001899D1AB1165 CRC64;

Query Match 5.7%; Score 101; DB 1; Length 394;
Best Local Similarity 21.5%; Pred. No. 27;
Matches 53; Conservative 43; Mismatches 66; Indels 84; Gaps 13;
QY 7 HGNLOYAEIPKSEPKVIEKAYIPVIETLIKEPFGNITGYTLKFLPKDIIILVKGGIASD 66
Db 72 HCN-----IVLALSPIIGLII-----GLGI-GYFLY-----TGGWASG 104
QY 67 LIEIIGTSYTAILPLPLSRVEAQVQRDRVKEELFEVSPKGFWLPPELADPIIPAILKDNG 126
Db 105 DVIILG-AYSALLPYVP-----DSAKY-----PPSYSLPM-----NA 137
QY 127 YEYLFADAMLSAHLNSAIKPIKPLPHLIKAQREKFRFYISYLLRLRELKAIKLVFEKG 186
Db 138 FTILFNSLLLIFFLILVSV-----VGLAVKGIKRELVGVFREG- 176
QY 187 VTLKVKDIEAVPVVAVNTAVMLIGRLPL-MNPKKVASWIEDKNILLYGTDTIEFIGYRDI 245
Db 177 ----IRNVVEVTLW--INFGAVLLAFVSLHMSIPKIISWILTFALILFSRKLVG--DV 228
QY 246 AGMSV 251
Db 229 LGVLSI 234

RESULT 10
O32450 PRELIMINARY; PRT; 653 AA.
AC O32450;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE 4-ALPHA-GLUCANOTRANSFERASE.
OS Pyrococcus kodakaraensis.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
OX NCBI_TaxID=69014;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KOD1;
RA Tachibana Y., Fujiwara S., Takagi M., Imanaka T.;
RL J. Ferment. Bioeng. 83:540-548(1997).
DR EMBL; D87907; BAA22062.1;
KW Transferase.
SQ SEQUENCE 653 AA; 76664 MW; B0C4695613F29219 CRC64;

Query Match 5.7%; Score 100.5; DB 1; Length 653;
Best Local Similarity 19.3%; Pred. No. 63;
Matches 74; Conservative 67; Mismatches 127; Indels 115; Gaps 19;
QY 24 EKAYIPVIETLIKEP---FGLNITGYTLKFL-----PKDIILVKGGIASDLIEIIGTS-Y 75
Db 28 ERSYRPFMETL-EEYPNMKVAVHYSGPILLEWIRDNKPEHDLRLSLVRKQLEIVVAGFY 86
QY 76 TAILPLPLSRVEAQVQRDRVKE--ELFEVSPKGFWLPE-LADPIIPAILKNDNGYEYLF 132
Db 87 EPVLASIP--KEDRIVQIEKLKEFARNLGYEARGVWLTQVWQPELVKSLRAAGIDYVIV 144
QY 133 DEAMLSAHLNS-----AIKPIKPLPHLIKAQREKFRFYISYLLRLREL 176
Db 145 DDYHMSAGLSKDELFWPYTDDGGEVITVFI-----DEKLRYL--IPFRPVD 191
QY 177 KAIKLVF-----EGKVTLVKVKDIEAVPVVAVNTAVMLIGRLPLMNPKKVASWIEDK-- 228
Db 192 KTLEYLHSLDDGDESKVAVFHDGGEKFGVWPG-----TYEWVYEKGW 233
QY 229 -----NILLYGTDIE-----FIGYRDIAG--RMSVEGL-----LEVID 259
Db 234 LREFFDRVSSDERINLMLYSEYLQFRPRGLVLYPIASYFEMSEWSLPAQAKLFVEFVE 293
QY 260 ELNSELCPSELKHSGRRELYLRTSSWADKSLRIWREDEGNARLNMLYNMRGELAF 319
Db 294 ELK-----KENKFDYRVFVGRGIWKNFFFKYKPESNYMHKRMMLVSKA-----VRNNPE 342
QY 320 ARGWLPERRLDAFRAIYNDWRG 342
Db 343 AREFILRAQCNDAY-----WHG 359

RESULT 11
Q9PKB1 PRELIMINARY; PRT; 690 AA.
ID Q9PKB1
AC Q9PKB1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL PROTEIN TC0556.
GN TC0556.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,

Db 118 REVNGKTIWIKDPLKQSEV-----EALYELPERGEVIARYKDG-KPAALIKINNKI 170
QY 129 YL--FADEAMLFSAHLNSAIKPIKPLPHLIKAQREKRFRYISYLLRLRELKAIKLVFEG 185
Db 171 YVGFPDEDVLANLIYIHIIVKKTNSPLPYIL-----FTVFLTLASL-----MLTFQ- 216
QY 186 KVTLLKVKDIEAVPVVAVNTAVMLIGRLPLPNPKKVASWIEDKNILLYGTDIEFIGYRDI 245
Db 217 -ETLKKKFLLELISALASVK--VFILSRINLLDEEKV-----LLNDRREIYNY--- 261
QY 246 AGRMSVEGLLEVIDELNSELCPSELKHSRGRELY-----LRTSSWADKSLRIWREDEGNA 299
Db 262 -----ILDNPGCHL-----RELKLNKPKVSTLTW---HLRILEK-----A 294
QY 300 RLNMLYNMRGELAFLAENSDARGWPLPERRLDAPRAIY 337
Db 295 NLIKSKKFGNRLIYYPADMDMRDLPLLYLKNETOKSIF 332

RESULT 14

Q9AVF4
ID Q9AVF4 PRELIMINARY; PRT; 608 AA.

AC Q9AVF4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MG-CHLATASE SUBUNIT CHLH (FRAGMENT).
GN MG-CH.

OS Amaranthus tricolor (Tampala).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Amaranthaceae; Amaranthus.
OX NCBI_TaxID=29722;
RN [1]

RP SEQUENCE FROM N.A.

RA Iwamoto K., Fukuda H., Sugiyama M.;
RT "Elimination of POR Expression correlates with Red Leaf Formation in
RT Amaranthus tricolor."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB050121; BAB41188.1; -
FT NON_TER 1
FT NON_TER 608
SQ SEQUENCE 608 AA; 67073 MW; 5D262E62BE2343CD CRC64;

Query Match

Best Local Similarity 5.6%; Score 99; DB 10; Length 608;
Matches 79; Conservative 62; Mismatches 115; Indels 104; Gaps 19;

QY 13 AEIPKSE-PKVIEWKAYIPVIETLIKEEPFGLNITGYTLKFLPKDIILV-----KGGI 63
Db 111 SELPTAERDLVVRKIYAKIMEIESRLPCGLHVIGEPPTALEAVATLVNIAALDRPDGI 170
QY 64 AS-----DLIEIIGTSYTAILPLPLSLRVEAQVQDRV-----KEELFE 102
Db 171 SSLPSILAQTVGRDIEDVYRGNDRGILKDVELLRQITETSRDAITAFVERTTNDKGQVNV 230
QY 103 VSPK-----GFWLPELADPIIPAILKNDNGEYLFADAEAML--FSAHLNSAIKPIKPLPHL 155
Db 231 VNDKLTSLTGLFGLNE-----PWIQYLSNTKFYRADREKLRVLFAFLGCECLK-----L 277
QY 156 IKAQREKRFRYISYLLRLRELKAI-----KLVEGKVTLLKVKDIEAVP 198
Db 278 VVADNE-----LGSLLKQALEGKYVEPGPGDPIRNPVKVLTGK-NIHALDPQAIP 326
QY 199 VVAVNTAVMLIGRLPLMNPCKV---ASWIEDKNILLYGTD-----IEFIGYR 243
Db 327 TTAAMQSAKVVVDR--LLERQKADGGKYPETVALVLWGTDNLIKTYGESLAQVLMMIGVR 384
QY 244 DIA---GRMS-VEGLLEVIDELNSELCPSELKHSRG--RELYLRTSSWADKSLRIWRE-DE 296
Db 385 PVADTFGRVKNKVEPV--SLEELGRPTVDVVVNCSGVFRDLFINQMNLDRAVKMLTELDE 442

RESULT 15

O17474
ID O17474 PRELIMINARY; PRT; 742 AA.

AC O17474;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CYCLIC GMP-DEPENDENT PROTEIN KINASE.
GN HYGK.

OS Hydra oligactis (Hydra).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydroidae; Hydra.
OX NCBI_TaxID=6088;
RN [1]

RP SEQUENCE FROM N.A.

RA Haeusler C.;

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF031931; AAC23588.1; -
DR HSSP; P05132; IATP.

DR InterPro; IPR000595; cNMP_binding.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000215; Serpin.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR InterPro; IPR000531; TonB_boxC.

DR Pfam; PF00027; cNMP_binding; 2.

DR Pfam; PF00069; pkinase; 1.

DR SMART; SM00100; cNMP; 2.

DR SMART; SM00220; S_TKC; 1.

DR SMART; SM00133; S_TK_X; 1.

DR PROSITE; PS00888; CNMP_BINDING_1; 1.

DR PROSITE; PS00889; CNMP_BINDING_2; 2.

DR PROSITE; PS00042; CNMP_BINDING_3; 2.

DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE; PS00284; SERPIN; UNKNOWN_1.

DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.

KW ATP-binding; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 742 AA; 85345 MW; B105D8C09618D154 CRC64;

Query Match

Best Local Similarity 5.6%; Score 99; DB 5; Length 742;
Matches 76; Conservative 60; Mismatches 133; Indels 82; Gaps 14;

QY 10 LQYAEIPKSEPKVIEWKAYIPVIETLIKEEPFGLNITGYTLKF--LPKDILVKG----- 61
Db 21 LDFARLKHLYPKL--RQDLKKKDELIEEK--NKKLSGYEIDILNLKKEIRQLQSVIEATR 76
QY 62 --GIASDLIEIIGTSYTAILPLPLSLRVEAQVQVOR-----DRVKEELFEVSPK-GFWLPEL 113
Db 77 NTNVDVIVVEEEDNGILLSRHLSQINHPKSRFIEVVKTVRNKRFAVSAESGKELQNL 136
QY 114 ADPIIP-----AILKNDNGEYLFADAEAMLFSAHLNSAIKPIKPLPHLIKAQREK 162
Db 137 DFPKIPKSTEVKEFTISQAVLKNFLKHL--EECQVKEIVLFMSQSKFKRGEYIIKEGDMG 194
QY 163 RFRYISYLLRLRELKAIKLVFEGKV-----TLKVKDIEAVPVW----- 200
Db 195 NALFVSYVGLLEISQDGKIL--GKPLRPGELFGEALILYNCTRTRASVKAIDDEVWYLER 252
QY 201 -----VAVNTAVM-----LIGRLPLMNPCKKVASWIEDKNILLYGTDIEFI-- 240
Db 253 NVFQAVMQKTCIMRREEHYNFLHSPVPKPNLPNDTLLKIVIVEIEE-----FYEDGEFIVR 308
QY 241 -GYRDIAGRMSVEGLLEVIDELNSELCPSELKHSRGRELYLRTSSWADKSLR 290

Db 309 EGERGDSFYILKQGLVKVLQMIIEGKDEPVEIRHLSQGEYFGEKALLGEDVR 359

Search completed: April 4, 2002, 14:44:28
Job time: 252 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 14:41:21 ; Search time 25.55 seconds
(without alignments)
1003.106 Million cell updates/sec

Title: US-09-407-806A-4
Perfect score: 346
Sequence: 1 LRALVFHGNLQYAEIPKSEP.....ERRLDAFRAIYNDRNGNEP 346

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 0

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 19 | 5.5 | 364 | AAW34643 | Thermotable alpha |
| 2 | 8 | 2.3 | 466 | AAW59165 | L. lactis MG1316 g |
| 3 | 8 | 2.3 | 624 | AAW35671 | Chlamydia pneumonia |
| 4 | 7 | 2.0 | 47 | AAG77652 | Human colon cancer |
| 5 | 7 | 2.0 | 50 | AAW30440 | Amino acid sequenc |
| 6 | 7 | 2.0 | 83 | AAW08113 | Arabidopsis thalia |
| 7 | 7 | 2.0 | 83 | AAW40493 | Arabidopsis thalia |
| 8 | 7 | 2.0 | 108 | AAW08112 | Arabidopsis thalia |
| 9 | 7 | 2.0 | 108 | AAW40492 | Arabidopsis thalia |
| 10 | 7 | 2.0 | 113 | AAW13155 | Human secreted pro |
| 11 | 7 | 2.0 | 116 | AAW12375 | Human 5' EST secre |

| | | | | | | |
|----|---|-----|-----|----|----------|--------------------|
| 12 | 7 | 2.0 | 138 | 21 | AAW54427 | Human pancreatic c |
| 13 | 7 | 2.0 | 140 | 20 | AAW37528 | Amino acid sequenc |
| 14 | 7 | 2.0 | 142 | 21 | AAG40491 | Arabidopsis thalia |
| 15 | 7 | 2.0 | 143 | 21 | AAG08111 | Arabidopsis thalia |
| 16 | 7 | 2.0 | 156 | 21 | AAG04247 | Arabidopsis thalia |
| 17 | 7 | 2.0 | 178 | 20 | AAW50041 | Buchnera aphidicol |
| 18 | 7 | 2.0 | 225 | 21 | AAG20264 | Arabidopsis thalia |
| 19 | 7 | 2.0 | 225 | 21 | AAG50345 | Arabidopsis thalia |
| 20 | 7 | 2.0 | 253 | 22 | AAU04821 | Micromonospora eve |
| 21 | 7 | 2.0 | 268 | 21 | AAG20263 | Arabidopsis thalia |
| 22 | 7 | 2.0 | 268 | 21 | AAG50344 | Arabidopsis thalia |
| 23 | 7 | 2.0 | 286 | 21 | AAW35951 | Human 3-hydroxyacy |
| 24 | 7 | 2.0 | 301 | 21 | AAW87599 | Human fatty acid b |
| 25 | 7 | 2.0 | 306 | 22 | AAG92081 | C glutamicum prote |
| 26 | 7 | 2.0 | 310 | 20 | AAW13370 | Amino acid sequenc |
| 27 | 7 | 2.0 | 310 | 21 | AAW27650 | Human protein PRO2 |
| 28 | 7 | 2.0 | 310 | 22 | AAW80238 | Human PRO238 prote |
| 29 | 7 | 2.0 | 325 | 21 | AAW92513 | Human OXRE-10. Ho |
| 30 | 7 | 2.0 | 325 | 21 | AAW73426 | Human secreted pro |
| 31 | 7 | 2.0 | 325 | 22 | AAW38738 | Human polypeptide |
| 32 | 7 | 2.0 | 327 | 22 | AAW40524 | Human polypeptide |
| 33 | 7 | 2.0 | 337 | 22 | AAW42410 | Human polypeptide |
| 34 | 7 | 2.0 | 350 | 19 | AAW74861 | Human secreted pro |
| 35 | 7 | 2.0 | 358 | 22 | AAG92635 | C glutamicum prote |
| 36 | 7 | 2.0 | 382 | 21 | AAG20262 | Arabidopsis thalia |
| 37 | 7 | 2.0 | 382 | 21 | AAG50343 | Arabidopsis thalia |
| 38 | 7 | 2.0 | 460 | 20 | AAW02536 | Human ICAM-6 prote |
| 39 | 7 | 2.0 | 471 | 22 | AAG98380 | Escherichia coli p |
| 40 | 7 | 2.0 | 484 | 18 | AAW21897 | Glutamyl-tRNA synt |
| 41 | 7 | 2.0 | 486 | 20 | AAW14129 | Human IL-1RD10 pro |
| 42 | 7 | 2.0 | 696 | 21 | AAW95298 | Human Xrec2. Homo |
| 43 | 7 | 2.0 | 867 | 20 | AAW37095 | Protein involved i |
| 44 | 7 | 2.0 | 870 | 20 | AAW35233 | C. pneumoniae prot |
| 45 | 7 | 2.0 | 999 | 21 | AAW73362 | HTRM clone 2674047 |

ALIGNMENTS

RESULT 1

AAW34643

ID AAW34643 standard; Protein; 364 AA.

XX

AC AAW34643;

XX

DT 27-MAR-1998 (first entry)

XX

DE Thermotable alpha-galactosidase AEDII12RA-alpha-gal-18GC.

XX

KW Alpha-galactosidase; alpha-glycosidase; thermostable enzyme;

KW food processing; alpha glycoside hydrolysis; raffinose;

KW stachyose; verbascose; bean; flatulence; AEDII12RA-alpha-gal-18GC.

XX

OS Thermococcus alcaliphilus strain AEDII12RA.

XX

FH Key Location/Qualifiers

FT Misc-difference 329

FT /note= "encoded by CTT"

XX

PN WO9732974-A1.

XX

PD 12-SEP-1997.

XX

PF 05-FEB-1997; 97WO-US01452.

XX

PR 08-MAR-1996; 96US-0613220.

XX

PA (RECO-) RECOMBINANT BIOCATALYSIS INC.

XX

PI Murphy D, Reid J, Rudolph MJ;

XX

DR WPI; 1997-470541/43.

DR N-PSDB; AAT93753.

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 PAILKDNG 126
|||||
Db 503 pailkdng 510

RESULT 4
AAG77652
ID AAG77652 standard; Protein; 47 AA.
XX
AC AAG77652;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:8418.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR N-PSDB; AAH37059.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 11; Page 9686; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 47 AA;

Query Match 2.0%; Score 7; DB 22; Length 47;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

-QY 165 RYISYLL 171
|||||
Db 24 ryisyll 30

RESULT 5
AAB30440
ID AAB30440 standard; peptide; 50 AA.
XX
AC AAB30440;
XX
DT 06-MAR-2001 (first entry)
XX
DE Amino acid sequence of a fragment of a galectin family member.
XX
KW Galectin; Galectin-3; cell growth; tumour; metastasis; cirrhotic liver;
KW hepatocellular carcinoma; liver neoplasm.
XX
OS Gallus sp.
XX
PN WO200062076-A1.
XX
PD 19-OCT-2000.
XX
PF 29-MAR-2000; 2000WO-US08561.
XX
PR 13-APR-1999; 99US-0129111.
XX
PA (HSUD/) HSU D K.
XX
PI Hsu DK, Liu F, Dowling CA;
XX
DR WPI; 2000-679520/66.
XX
PT Cirrhotic liver and hepatocellular carcinoma marker galectin-3, useful
PT in diagnostic, prognostic and therapeutic methods for treating and
PT preventing hepatocellular carcinoma -
XX
PS Example 3; Page 35; 49pp; English.
XX
CC AAB30440-85 represent peptides derived from galectin family members.
CC Galectin-3 has been associated with assorted processes such as cell
CC growth, tumour transformation, and metastasis. It is composed of two
CC domains: a carbohydrate-binding region on the carboxy terminal which
CC binds saccharides with terminal galactose residues, and an amino terminal
CC domain consisting of tandem repeats of nine amino acids. The expression
CC of galectin-3 is induced in cirrhotic liver and hepatocellular
CC carcinoma, and so is useful as a marker of cirrhotic liver and
CC hepatocellular carcinoma. Galectin-3 is used in diagnostic, prognostic
CC and therapeutics for the treatment and prevention of hepatocellular
CC carcinoma. It is useful for determining whether an individual is
CC suffering from hepatocellular carcinoma or is likely to suffer from
CC hepatocellular carcinoma or any neoplasm of liver origin or cirrhosis of
CC the liver in the future.
XX
SQ Sequence 50 AA;

Query Match 2.0%; Score 7; DB 21; Length 50;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 LPLLPLS 85
|||||
Db 15 lpllpls 21

RESULT 6
AAG08113
ID AAG08113 standard; Protein; 83 AA.
XX
AC AAG08113;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 5523.

| | | | |
|----|--|--------------|---------------|
| XX | | | |
| KW | Protein identification; signal transduction pathway; metabolic pathway; | 23-JUN-1999; | 99US-0140353. |
| KW | hybridisation assay; genetic mapping; gene expression control; promoter; | 23-JUN-1999; | 99US-0140354. |
| KW | termination sequence. | 24-JUN-1999; | 99US-0140695. |
| XX | | 28-JUN-1999; | 99US-0140823. |
| OS | Arabidopsis thaliana. | 29-JUN-1999; | 99US-0140991. |
| XX | | 30-JUN-1999; | 99US-0141287. |
| PN | EP1033405-A2. | 01-JUL-1999; | 99US-0141842. |
| XX | | 01-JUL-1999; | 99US-0142154. |
| PD | 06-SEP-2000. | 02-JUL-1999; | 99US-0142055. |
| XX | | 06-JUL-1999; | 99US-0142390. |
| PF | 25-FEB-2000; 2000EP-0301439. | 08-JUL-1999; | 99US-0142803. |
| XX | | 09-JUL-1999; | 99US-0142920. |
| PR | 25-FEB-1999; 99US-0121825. | 12-JUL-1999; | 99US-0142977. |
| PR | 05-MAR-1999; 99US-0123180. | 13-JUL-1999; | 99US-0143542. |
| PR | 09-MAR-1999; 99US-0123548. | 14-JUL-1999; | 99US-0143624. |
| PR | 23-MAR-1999; 99US-0125788. | 15-JUL-1999; | 99US-0144005. |
| PR | 25-MAR-1999; 99US-0126264. | 16-JUL-1999; | 99US-0144085. |
| PR | 29-MAR-1999; 99US-0126785. | 16-JUL-1999; | 99US-0144086. |
| PR | 01-APR-1999; 99US-0127462. | 19-JUL-1999; | 99US-0144325. |
| PR | 06-APR-1999; 99US-0128234. | 19-JUL-1999; | 99US-0144331. |
| PR | 08-APR-1999; 99US-0128714. | 19-JUL-1999; | 99US-0144332. |
| PR | 16-APR-1999; 99US-0129845. | 19-JUL-1999; | 99US-0144333. |
| PR | 19-APR-1999; 99US-0130077. | 19-JUL-1999; | 99US-0144334. |
| PR | 21-APR-1999; 99US-0130449. | 19-JUL-1999; | 99US-0144335. |
| PR | 23-APR-1999; 99US-0130510. | 20-JUL-1999; | 99US-0144352. |
| PR | 23-APR-1999; 99US-0130891. | 20-JUL-1999; | 99US-0144632. |
| PR | 28-APR-1999; 99US-0131449. | 20-JUL-1999; | 99US-0144884. |
| PR | 30-APR-1999; 99US-0132048. | 21-JUL-1999; | 99US-0144814. |
| PR | 04-APR-1999; 99US-0132407. | 21-JUL-1999; | 99US-0145086. |
| PR | 05-MAY-1999; 99US-0132484. | 21-JUL-1999; | 99US-0145088. |
| PR | 06-MAY-1999; 99US-0132485. | 22-JUL-1999; | 99US-0145085. |
| PR | 06-MAY-1999; 99US-0132486. | 22-JUL-1999; | 99US-0145087. |
| PR | 06-MAY-1999; 99US-0132487. | 22-JUL-1999; | 99US-0145089. |
| PR | 07-MAY-1999; 99US-0132863. | 22-JUL-1999; | 99US-0145192. |
| PR | 11-MAY-1999; 99US-0134256. | 23-JUL-1999; | 99US-0145145. |
| PR | 14-MAY-1999; 99US-0134218. | 23-JUL-1999; | 99US-0145218. |
| PR | 14-MAY-1999; 99US-0134219. | 23-JUL-1999; | 99US-0145224. |
| PR | 14-MAY-1999; 99US-0134221. | 26-JUL-1999; | 99US-0145276. |
| PR | 14-MAY-1999; 99US-0134370. | 27-JUL-1999; | 99US-0145913. |
| PR | 18-MAY-1999; 99US-0134768. | 27-JUL-1999; | 99US-0145918. |
| PR | 19-MAY-1999; 99US-0134941. | 27-JUL-1999; | 99US-0145919. |
| PR | 20-MAY-1999; 99US-0135124. | 28-JUL-1999; | 99US-0145951. |
| PR | 21-MAY-1999; 99US-0135353. | 02-AUG-1999; | 99US-0146386. |
| PR | 24-MAY-1999; 99US-0135629. | 02-AUG-1999; | 99US-0146388. |
| PR | 25-MAY-1999; 99US-0136021. | 02-AUG-1999; | 99US-0146389. |
| PR | 27-MAY-1999; 99US-0136392. | 03-AUG-1999; | 99US-0147038. |
| PR | 28-MAY-1999; 99US-0136782. | 04-AUG-1999; | 99US-0147204. |
| PR | 01-JUN-1999; 99US-0137222. | 04-AUG-1999; | 99US-0147302. |
| PR | 03-JUN-1999; 99US-0137528. | 05-AUG-1999; | 99US-0147192. |
| PR | 04-JUN-1999; 99US-0137502. | 05-AUG-1999; | 99US-0147260. |
| PR | 07-JUN-1999; 99US-0137724. | 06-AUG-1999; | 99US-0147303. |
| PR | 08-JUN-1999; 99US-0138094. | 06-AUG-1999; | 99US-0147416. |
| PR | 10-JUN-1999; 99US-0138540. | 09-AUG-1999; | 99US-0147493. |
| PR | 10-JUN-1999; 99US-0138847. | 09-AUG-1999; | 99US-0147935. |
| PR | 14-JUN-1999; 99US-0139119. | 10-AUG-1999; | 99US-0148171. |
| PR | 16-JUN-1999; 99US-0139452. | 11-AUG-1999; | 99US-0148319. |
| PR | 16-JUN-1999; 99US-0139453. | 12-AUG-1999; | 99US-0148341. |
| PR | 17-JUN-1999; 99US-0139492. | 13-AUG-1999; | 99US-0148565. |
| PR | 18-JUN-1999; 99US-0139454. | 13-AUG-1999; | 99US-0148684. |
| PR | 18-JUN-1999; 99US-0139455. | 16-AUG-1999; | 99US-0149368. |
| PR | 18-JUN-1999; 99US-0139456. | 17-AUG-1999; | 99US-0149175. |
| PR | 18-JUN-1999; 99US-0139457. | 18-AUG-1999; | 99US-0149426. |
| PR | 18-JUN-1999; 99US-0139458. | 20-AUG-1999; | 99US-0149722. |
| PR | 18-JUN-1999; 99US-0139459. | 20-AUG-1999; | 99US-0149723. |
| PR | 18-JUN-1999; 99US-0139460. | 20-AUG-1999; | 99US-0149929. |
| PR | 18-JUN-1999; 99US-0139461. | 23-AUG-1999; | 99US-0149902. |
| PR | 18-JUN-1999; 99US-0139462. | 23-AUG-1999; | 99US-0149930. |
| PR | 18-JUN-1999; 99US-0139463. | 25-AUG-1999; | 99US-0150566. |
| PR | 18-JUN-1999; 99US-0139750. | 26-AUG-1999; | 99US-0150884. |
| PR | 18-JUN-1999; 99US-0139763. | 27-AUG-1999; | 99US-0151065. |
| PR | 21-JUN-1999; 99US-0139817. | 27-AUG-1999; | 99US-0151066. |
| PR | 22-JUN-1999; 99US-0139899. | 27-AUG-1999; | 99US-0151080. |
| | | 30-AUG-1999; | 99US-0151303. |

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 7; DB 21; Length 83;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 186 KVTCLKVK 192
Db 2 kvtlkvk 8
|||||

RESULT 7
AAG40493
ID AAG40493 standard; Protein; 83 AA.
XX
AC AAG40493;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 50250.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 28-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 7; DB 21; Length 83;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 KVTLLKVK 192
 |||||
 Db 2 kvtlkvk 8

RESULT 8
 AAG08112
 ID AAG08112 standard; Protein; 108 AA.
 XX
 AC AAG08112;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 5522.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 7; DB 21; Length 108;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 186 KVTLLKVK 192
Db 27 kvtlkvk 33
|||||

RESULT 9
AAG40492
ID AAG40492 standard; Protein; 108 AA.
XX
AC AAG40492;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 50249.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.

PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 7; DB 21; Length 108;
Best Local Similarity 100.0%; Pred.No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 KVTCLKVK 192
Db 27 kvtlkvk 33

RESULT 10

AAY13155

ID AAY13155 standard; Protein; 113 AA.

XX AAY13155;

XX 22-JUN-1999 (first entry)

XX Human secreted protein encoded by 5' EST SEQ ID NO: 169.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.

XX Homo sapiens.

XX WO9906552-A2.

XX 11-FEB-1999.

XX 31-JUL-1998; 98WO-IB01236.

XX 01-AUG-1997; 97US-0905223.

XX (GEST) GENSET.

XX Duclert A, Dumas Milne Edwards J, Lacroix B;

XX WPI; 1999-153782/13.

XX N-PSDB; AAX51955.

XX New isolated brain-derived nucleic acids - used to develop products
PT which may have cytokine, immune, regulatory, haematopoiesis
PT regulating, anti-inflammatory or tumour inhibition activity

XX PS Claim 34; Page 541; 577pp; English.

XX CC AAX51787 to AAX52019 represent 5' expressed sequence tags (ESTs) for

CC human secreted proteins, and encode the proteins given in AAY12987 to

CC AAY13219, respectively. The proteins given represent the signal peptide

CC and an N-terminal fragment of a secreted protein. The nucleic acid

CC sequences can be used for producing secreted human gene products. They

CC can also be used to develop products for diagnosis and therapy. The

CC proteins obtained may have cytokine activity, cell

CC proliferation/differentiation activity, haematopoiesis regulating

CC activity, tissue growth regulating activity, reproductive hormone

CC regulating activity, chemotactic/chemokinetic activity, haemostatic and

CC thrombolytic activity, receptor/ligand activity, anti-inflammatory

CC activity, tumour inhibition activity or other activities. The products

CC can be used in forensic, gene therapy and chromosome mapping procedures.

CC The sequences can also be used for obtaining corresponding promoter

CC sequences. The nucleic acids encoding the signal peptide can be used for

CC directing extracellular secretion of a polypeptide or the insertion of a

CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX SQ Sequence 113 AA;

Query Match 2.0%; Score 7; DB 20; Length 113;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 TAILPLL 82

Db 27 tailpll 33

RESULT 11

AAY12375

ID AAY12375 standard; Protein; 116 AA.

AC AAY12375;

17-JUN-1999 (first entry)

Human 5' EST secreted protein SEQ ID NO:406.

Human; secreted protein; EST; expressed sequence tag; diagnosis;

forensic; gene therapy; chromosome mapping; signal peptide;

upstream regulatory sequence; cytokine activity; cell proliferation;

differentiation; haematopoiesis regulation; tissue growth regulation;

reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;

thrombolytic; anti-inflammatory; tumour inhibition.

Homo sapiens.

WO9906548-A2.

11-FEB-1999.

31-JUL-1998; 98WO-IB01222.

01-AUG-1997; 97US-0905135.

(GEST) GENSET.

Duclert A, Dumas Milne Edwards J, Lacroix B;

WPI; 1999-153778/13.

N-PSDB; AAX41208.

New nucleic acids encoding human secreted proteins - obtained from

cDNA libraries prepared from e.g. liver, ovary, brain, prostate,

kidney, lung, umbilical cord, placenta and colon tissue

Claim 27; Page 727-728; 824pp; English.

CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for

CC human secreted proteins, and encode the proteins given in AAY12261 to

CC AAY12514, respectively. The proteins given represent the signal peptide

CC and an N-terminal fragment of a secreted protein. The nucleic acid

CC sequences can be used for producing secreted human gene products. They

CC can also be used to develop products for diagnosis and therapy. The

CC proteins obtained may have cytokine activity, cell

CC proliferation/differentiation activity, haematopoiesis regulating

CC activity, tissue growth regulating activity, reproductive hormone

CC regulating activity, chemotactic/chemokinetic activity, haemostatic and

CC thrombolytic activity, receptor/ligand activity, anti-inflammatory

CC activity, tumour inhibition activity or other activities. The products

CC can be used in forensic, gene therapy and chromosome mapping procedures.

CC The sequences can also be used for obtaining corresponding promoter

CC sequences. The nucleic acids encoding the signal peptide can be used for

CC directing extracellular secretion of a polypeptide or the insertion of a

CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX SQ Sequence 116 AA;

Query Match 2.0%; Score 7; DB 20; Length 116;

Best Local Similarity 100.0%; Pred. No. 52;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 TAILPLL 82

Db 7 tailpll 13

RESULT 12

AAB54427

ID AAB54427 standard; Protein; 138 AA.

AC AAB54427;

09-MAR-2001 (first entry)

Human pancreatic cancer antigen protein sequence SEQ ID NO:879.

Human; pancreas; pancreatic cancer; pancreatic cancer antigen;

detection; diagnosis; identification; cytostatic; neuroprotective;

nootropic; immunomodulatory; relaxant; contraceptive; gynaecological;

antiinflammatory; cardiac; gene therapy; chromosome mapping;

linkage analysis; tissue identification; tissue typing; forensic;

neural; immune system; muscular; reproductive; gastrointestinal;

pulmonary; cardiovascular; renal; proliferative.

Homo sapiens.

WO200055320-A1.

21-SEP-2000.

08-MAR-2000; 2000WO-US05989.

12-MAR-1999; 99US-0124270.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI; 2000-579444/54.

N-PSDB; AAC99192.

New nucleic acid that is a pancreatic cancer antigen for preventing,

treating, or ameliorating a medical condition, particular pancreatic

cancer, or for use in assays for diagnosing a pathological condition -

Claim 11; Page 1339; 1379pp; English.

AAC98773 to AAC99231 encode the human pancreatic cancer associated

proteins, called pancreatic cancer antigens, given in AAB54008 to

CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.

SQ Sequence 138 AA;

Query Match 2.0%; Score 7; DB 21; Length 138;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 RGEIAFL 314
| | | | |
Db 4 rgeiafl 10

RESULT 13
AAY37528
ID AAY37528 standard; Protein; 140 AA.

XX AAY37528;

XX 07-OCT-1999 (first entry)

XX Amino acid sequence of a Chlamydia trachomatis protein.

KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.

XX Chlamydia trachomatis.

XX WO9928475-A2.

XX 10-JUN-1999.

XX 27-NOV-1998; 98WO-IB01939.

XX 04-NOV-1998; 98US-0107077.

XX 28-NOV-1997; 97FR-0015041.

XX 17-DEC-1997; 97FR-0016034.

XX (GEST) GENSET.

XX Griffais R;

XX WPI; 1999-371125/31.

XX Genome sequence of Chlamydia trachomatis

XX Disclosure; Page 1196; 1755pp; English.

CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences

CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
CC perihepatitis, bartholinitis; pneumopathy; venereal lymphogranulomatosis;
CC and venereal lymphogranulomatosis. The polypeptides of the invention
CC may be of use in treating these diseases.

SQ Sequence 140 AA;

Query Match 2.0%; Score 7; DB 20; Length 140;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 KEELFEV 103
| | | | |
Db 102 keelfev 108

RESULT 14
AAG40491

ID AAG40491 standard; Protein; 142 AA.

XX AAG40491;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 50248.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132486.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 7; DB 21; Length 142;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 KVTLLKVK 192
| | | | | | | |
Db 61 kvtlkvk 67

RESULT 15
AAG08111
ID AAG08111 standard; Protein; 143 AA.

XX AC AAG08111;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 5521.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 7; DB 21; Length 143;
Best Local Similarity 100.0%; Pred. No. 63;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 186 KVTCLKVK 192
Db 62 kvtlkvk 68

Search completed: April 4, 2002, 14:45:30
Job time: 249 sec

